



مهندسی کامپیوتر  
گروه شبکه‌های کامپیوتری

# مروری بر سازوکارهای گسترش ژرف برای سامانه‌های ارتباطات بی‌سیم نسل بعد

گزارش سمینار کارشناسی ارشد

در رشته شبکه‌های کامپیوتری

علی اصغر رجبی

استاد راهنما:

دکتر وصال حکمی

آبان ۱۴۰۴

## چکیده

شبکه‌های سیار نسل آینده از جمله نسل ششم، با رشد تصاعدی تقاضا برای منابع ناشی از افزایش دستگاه‌های متصل، مواجه هستند. نیازهای کلیدی نسل ششم شامل نرخ داده بالا (یک ترابیت بر ثانیه)، بازده طیفی بالا، اتصال انبوه مانند دستگاه‌های اینترنت اشیا که ده برابر نسل پنجم هستند و تأخیر کم است. همچنین ظهور کاربردهای مبتنی بر هوش مصنوعی نیازمند یکپارچه‌سازی هوش مصنوعی و ارتباطات است [۱، ۲]. روش‌های سنتی ارتباطات بی‌سیم، که اغلب بر اصول بهینه‌سازی تکرارشونده و تعامد منابع تکیه دارند، اغلب به دلیل پیچیدگی بالا و زمان اجرای طولانی که اغلب ناشی از متغیر بودن پارامترهای محیطی است برای برآوردن الزامات بی‌سابقه سناریوهای بلادرنگ در شبکه‌های نسل جدید ناکافی هستند [۱، ۲].

بسیاری از مسائل حیاتی، مانند تخصیص منابع، غیرمحدب هستند و یافتن راه حل بهینه سراسری برای آن‌ها دشوار است [۳، ۴، ۵، ۶]. گسترش ژرف به عنوان یک رویکرد یادگیری ژرف مدل-محور مطرح می‌شود [۱]. این سازوکار به منظور بهره‌گیری همزمان از مزایای دانش دامنه الگوریتم‌های بهینه‌سازی مرسوم و کارایی محاسباتی مدل‌های یادگیری ژرف توسعه یافته است [۱، ۲، ۶]. در این سمینار، ضمن معرفی حوزه پژوهشی «گسترش ژرف» بررسی می‌شود که چگونه می‌توان راه‌حلی تفسیرپذیر، مقیاس‌پذیر و بسیار کارآمد برای چالش‌های لایه فیزیکی و تخصیص منابع در شبکه‌های پیچیده و توزیع شده نسل آینده ارائه داد [۱، ۶].

**کلمات کلیدی:** گسترش ژرف، یادگیری ژرف، یادگیری ماشین، مدل-محور، شبکه‌های ارتباطی بی‌سیم، پردازش سیگنال، لایه فیزیکی، تخصیص منابع

# فهرست مطالب

۱	فصل ۱	مقدمه
۳	۱.۱	اهمیت و ضرورت
۴	۲.۱	چالش‌ها
۵	۳.۱	ساختار گزارش
۶	فصل ۲	مفاهیم پایه‌ای
۱۰	فصل ۳	مروری بر کارهای پیشین
۱۱	۱.۳	تخمین کانال
۱۱	۲.۳	آشکارسازی کانال
۱۱	۳.۳	طراحی پیش‌کدگذار
۱۱	۴.۳	سنجش و ارتباطات
۱۱	۵.۳	رمزگشایی در کدهای تصحیح خطا
۱۱	۶.۳	تخصیص توان
۲۴	فصل ۴	نتیجه‌گیری و کارهای آینده
۲۴	۱.۴	نتیجه‌گیری
۲۴	۲.۴	کارهای آینده
۲۵		مراجع
۲۶		واژه نامه انگلیسی به فارسی
۲۹		واژه نامه فارسی به انگلیسی
۳۲		نمایه







# فهرست جدول‌ها





## Numbers

$\Delta G$  ..... Fifth Generation

## A

ADMM ..... Alternating Direction Method of Multipliers

AR ..... Augmented Reality

## B

BnB ..... Branch and Bound

BP ..... Belief Propagation

## C

CD-NOMA ..... Code Domain Non Orthogonal Multiple Access

CNN ..... Convolutional Neural Network

## D

DRL ..... Deep Reinforcement Learning

## **F**

FUM ..... Full-Unfolding Based Model

## **G**

GAN ..... Generative Adversarial Network

GNN ..... Graph Neural Network

## **I**

IoT ..... Internet of Things

ISAC ..... Integrated Sensing and Communication

ISTA ..... Iterative Shrinkage-Thresholding Algorithm

## **J**

JCEHB ..... Joint Channel Estimation and Hybrid Beamforming

## **K**

KKT ..... Karush-Kuhn-Tucker

## **L**

LDPC ..... Low-Density Parity-Check

LS ..... Least Square

## M

MASUM .....	Multi-Attention-based Semi-Unfolding Model
MIMO .....	Multiple Input Multiple Output
MLP .....	Multilayer Perceptron
MM .....	Multimodel
mmWave .....	Millimeter Wave
MRA .....	Multiple Random Access

## N

NGMA .....	Next Generation Multiple Access
NOMA .....	Non Orthogonal Multiple Access

## O

ORAN .....	Open Radio Access Network
------------	---------------------------

## P

PD-NOMA .....	Power Domain Non Orthogonal Multiple Access
PGA .....	Parallel Genetic Algorithm
PGD .....	Projected Gradient Decent

## Q

QoS .....	Quality of Service
-----------	--------------------

## **R**

RNN ..... Recurrent Neural Network

## **S**

SCA ..... Successive Convex Approximation

SIC ..... Successive Interference Cancellation

SNR ..... Signal to Noise Ratio

SoRP ..... Sum-of-Ratios Problem

STARS ..... Simultaneously Transmitting And Reflecting Surface

## **T**

T&R NOMA ..... Transmitting and Reflecting Non-Orthogonal Multiple Access

## **U**

UE ..... User Equipment

## **V**

VR ..... Virtual Reality

## **W**

WMMSE ..... Weighted Minimum Mean Square Error

# فصل ۱

## مقدمه

شبکه‌های بی‌سیم نسل ششم در حال حاضر به عنوان چشم‌انداز آینده ارتباطات مطرح هستند و پیش‌بینی می‌شود که ویژگی‌هایی مانند پوشش گسترده منطقه‌ای، ارائه خدمات متنوع برای سناریوهای پیش‌بینی شده، اتصالات انبوه و ناهمگونی پویا را به همراه داشته باشند. این ویژگی‌های پیشرفته منجر به ایجاد مسائل بهینه‌سازی شبکه در مقیاس بزرگ و پیچیده‌ای می‌شوند. در دهه‌های اخیر، سامانه‌های ارتباطات بی‌سیم ستون فقرات جامعه مدرن بوده‌اند و کاربردهایی از ارتباطات شخصی تا اتوماسیون صنعتی را ممکن ساخته‌اند. با این حال، با رسیدن شبکه‌های نسل پنجم (5G<sup>1</sup>) به بلوغ تجاری، تقاضای فزاینده‌ای برای نرخ داده‌های بسیار بالا و برنامه‌های جدید به وجود آمده است که نیازمند اتصال انبوه هستند (مانند اینترنت اشیا (IoT<sup>2</sup>) و فراجهان<sup>۳</sup>). این افزایش تقاضا منجر به رقابت فزاینده‌ای برای تخصیص کارآمد طیف می‌شود؛ زیرا شبکه‌های آینده باید حجم عظیمی از داده‌ها را پشتیبانی کنند. روش‌های مرسوم مبتنی بر مدل با وجود کارایی در سناریوهای ساده که مدل‌های ریاضی دقیقی برای آن‌ها وجود دارد، در مواجهه با کاربردهای پیچیده و واقعی شبکه‌های نسل ششم، با چالش بار محاسباتی بالا و زمان پردازش طولانی دست و پنجه نرم می‌کنند. به طور موازی، پیشرفت‌های اخیر در یادگیری ماشین و به طور خاص یادگیری ژرف<sup>۴</sup>، رویکردهای نویدبخشی را برای حل مسائل پیچیده و غیرقابل حل در گذشته را در لایه فیزیکی و مدیریت منابع شبکه ارائه کرده‌اند. روش‌های یادگیری ژرف صرفاً مبتنی بر داده توانایی‌های تقریب قدرتمندی را ارائه می‌دهند و می‌توانند استنتاج برخط سریعی داشته باشند. با این حال، محدودیت‌های

<sup>1</sup>Fifth Generation  
<sup>2</sup>Internet of Things

<sup>3</sup>Metaverse  
<sup>4</sup>Deep Learning

این روش‌ها شامل نیاز به مجموعه داده‌های ناکافی یا بزرگ و همچنین ضعف در تفسیر پذیری<sup>۵</sup> است. در حوزه هوش مصنوعی، پژوهشگران به دنبال آن هستند که رایانه‌ها بتوانند دانش را مانند انسان درک کنند، به دست آورند، پردازش کنند و استنتاج نمایند. برای این منظور، مهندسی مبتنی بر دانش پیشنهاد شده است، که در آن سامانه‌های خبره برای تقلید از توانایی استدلال انسان ساخته می‌شوند. این مسائل، لزوم حرکت به سمت رویکردهای ترکیبی را نشان می‌دهد که بتوانند از مزایای ساختاریافته مدل‌های ریاضی و انعطاف‌پذیری و سرعت یادگیری ژرف، به طور همزمان استفاده کنند.

برای رفع چالش‌هایی که مدل‌های مرسوم و روش‌های صرفاً مبتنی بر داده با آن مواجه هستند، یادگیری ژرف دانش‌محور<sup>۶</sup> معرفی شده است. این رویکرد دانش دامنه<sup>۷</sup> را در شبکه‌های عصبی ادغام می‌کند و نقاط قوت هر دو روش مبتنی بر مدل و مبتنی بر داده را ترکیب می‌نماید. یادگیری ژرف دانش‌محور با هدف ارائه یک راهنمای روشن‌گر برای ترکیب موثر دانش دامنه در شبکه‌های عصبی در حوزه ارتباطات بی‌سیم، به دنبال پیشبرد شبکه‌های نسل ششم هوشمند، کارآمد و قابل اعتماد است. در این زمینه، مفهوم دانش دامنه خاص ارتباطات به طور جامع تعریف شده است؛ این دانش شامل نظریه‌های فنی و شناخت تجربی است که توسط دانشمندان و متخصصان در فرآیند تصور، ساخت و بهینه‌سازی شبکه‌های ارتباطی بی‌سیم انباشته شده است. این دانش به دو جنبه کلیدی تقسیم می‌شود. نخست، دانش در طول فرآیند مدل‌سازی مانند مدل‌سازی نظریه‌های وظایف بی‌سیم و دیگری، دانش در طول فرآیند تصمیم‌گیری مانند الگوریتم‌های نظری مبتنی بر مدل. به طور کلی، دانش دامنه در شبکه‌های بی‌سیم به دانش علمی و دانش تخصصی طبقه‌بندی می‌شود که اولی رسمی‌تر و دومین غیررسمی‌تر است. دانش علمی شامل قوانین نظری انتقال، روش‌های مدل‌سازی شبکه‌های بی‌سیم و راه‌حل‌های نظری بهینه‌سازی شبکه است، که اغلب به صورت رابطه‌های ریاضی رسمی بیان می‌شوند.

یکی از قدرتمندترین نمودهای یادگیری ژرف دانش‌محور، روش گسترش ژرف<sup>۸</sup> است که به طور گسترده در پردازش سیگنال و بهینه‌سازی شبکه‌های بی‌سیم نسل جدید مورد بررسی قرار گرفته است. این روش، الگوریتم‌های تکرارشونده مرسوم مانند روش‌های بهینه‌سازی یا تخمین سیگنال را به صورت لایه‌های ساختاریافته در یک شبکه عصبی ژرف<sup>۹</sup> نگاشت می‌کند. این رویکرد، قابلیت تفسیر تفسیر پذیری را با حفظ ساختار اصلی الگوریتم‌های تکرارشونده افزایش می‌دهد، در حالی که پارامترهای الگوریتم‌های مرسوم مانند اندازه‌های گام یا ضرایب همگرایی به پارامترهای قابل آموزش<sup>۱۰</sup> در شبکه عصبی ژرف تبدیل می‌شوند. این امر امکان استفاده از بهینه‌سازی مبتنی بر داده را فراهم می‌سازد تا همگرایی تسریع شود و عملکرد بهبود یابد. برخلاف قابلیت تفسیر پسینی<sup>۱۱</sup> که پس از آموزش مدل‌ها اتفاق می‌افتد، یادگیری ژرف دانش‌محور به نظام فکری قابلیت تفسیر پیشینی<sup>۱۲</sup> تعلق دارد که در آن تفسیرپذیری از همان ابتدا در طراحی مدل یا فرآیند یادگیری جاسازی می‌شود. این ویژگی به ویژه در شبکه‌های بی‌سیم که الزامات حیاتی ایمنی دارند مهم است.

پژوهش‌های انجام شده در منابع نشان می‌دهد که فقدان یک طبقه‌بندی یکپارچه و نظام‌مند از رویکردهای ادغام دانش برای بهینه‌سازی شبکه‌های بی‌سیم، یک شکاف تحقیقاتی مهم بوده است. در پاسخ به این نیاز، یک طبقه‌بندی جدید برای رویکردهای ادغام دانش در شبکه‌های بی‌سیم پیشنهاد شده است که بر اساس این موضوع طراحی شده که دانش دامنه در کدام اجزای خط

لوله<sup>۱۳</sup> یادگیری ژرف و به چه شکلی ادغام می‌شود. این طبقه‌بندی شامل ادغام دانش دامنه در انتخاب مدل شبکه عصبی، سفرای سازی

<sup>5</sup>Interpretability

<sup>6</sup>Knowledge-Driven Deep Learning

<sup>7</sup>Domain Knowledge

<sup>8</sup>Deep Unfolding

<sup>9</sup>Deep Neural Network

<sup>10</sup>Trainable Parameter

<sup>11</sup>post-hoc explainability

<sup>12</sup>ante-hoc explainability

<sup>13</sup>Pipeline

مدل شبکه عصبی، ساخت معماری ترکیب دانش و داده، طراحی تابع زیان<sup>۱۴</sup> و پیکربندی فرارامتر<sup>۱۵</sup> است. این چارچوب ساختاریافته، یک دستورالعمل عمل برای ادغام دانش و شبکه‌های عصبی ارائه می‌دهد. از منظر کاربرد، تمرکز اصلی رویکردهای دانش‌محور و گسترش ژرف در شبکه‌های بی‌سیم بر دو حوزه اساسی تخصیص منابع و پردازش سیگنال است. در حوزه تخصیص منابع، روش‌هایی مانند گسترش الگوریتم WMMSE<sup>۱۶</sup> با استفاده از شبکه‌های عصبی گرافی<sup>۱۷</sup> برای تخصیص توان کارآمد و مقیاس‌پذیر مورد استفاده قرار گرفته‌اند. در پردازش سیگنال نیز، شبکه عصبی گرافی به طور گسترده‌ای برای وظایفی نظیر تشخیص سیگنال، تخمین کانال و طراحی پیش‌کدگذارها پیش‌کدگذارها<sup>۱۸</sup> به کار گرفته شده است. این روش‌ها به ویژه در مواجهه با چالش‌های فناوری‌های نسل جدید مانند آرایه‌های بزرگ<sup>۱۹</sup> MIMO، امواج میلی‌متری<sup>۲۰</sup> و سطوح بازتابنده هوشمند<sup>۲۱</sup> اهمیت می‌یابند. برای مثال، سفارشی‌سازی مدل شبکه عصبی می‌تواند شامل طراحی ساختار زیربخش، طراحی ساختار کامل یا شبکه‌های عصبی ترکیبی ساختارمحور باشد. طراحی ساختار کامل شبکه‌های عصبی به ویژه از طریق رویکرد نوآورانه گسترش الگوریتم<sup>۲۲</sup> محقق می‌شود که ساختار بنیادی الگوریتم‌های تکرارشونده نظری را حفظ کرده و قابلیت تفسیر و استنتاج بر خط سریع را فراهم می‌سازد.

## ۱.۱ اهمیت و ضرورت

همانطور که گفته شد در نسل‌های آتی شبکه‌های تلفن همراه با رشد عظیم دستگاه‌های متصل و انتقال حجم بالایی از داده‌ها روبه‌رو هستند؛ بنابراین، هدف اصلی در نسل ششم دستیابی به طرح‌هایی است که فراتر از متعادل بودن منابع عمل کنند و عملکرد، قابلیت اطمینان و کارایی بالاتری را ارائه دهند. در این شبکه‌ها، کارایی انرژی<sup>۲۳</sup> یک حوزه حیاتی است، زیرا تقاضای فزاینده برای داده و تأکید جهانی بر کاهش مصرف انرژی این موضوع را به یک دغدغه اصلی تبدیل کرده است. مدیریت بهینه منابع بی‌سیم، مانند تخصیص توان و طیف، به منظور به حداکثر رساندن سودمندی‌های سطح سامانه نظیر نرخ جمعیتی وزن‌دار یا کارایی انرژی، در حالی که محدودیت‌های منابع و سخت‌افزاری رعایت شوند، یک چالش پیچیده محسوب می‌شود.

بسیاری از این مسائل بهینه‌سازی منابع، مانند مسائل تخصیص توان، غالباً غیرمحدب<sup>۲۴</sup> یا از نوع مسئله مجموع-نسبت‌ها SoRP<sup>۲۵</sup> هستند و به دست آوردن راه‌حل‌های بهینه سراسری برای آن‌ها در زمان چندجمله‌ای دشوار است. روش‌های مرسوم مدل‌محور، که بر اساس اصول ریاضی و نظریه ارتباطات طراحی شده‌اند، اغلب بر فرضیات ساده‌سازی شده و ایده‌آل درباره سامانه‌های ارتباطی، مانند فرض اختلال گاوسی و رفتارهای سامانه خطی، تکیه دارند؛ که این فرضیات معمولاً منجر به مدل‌های سامانه غیردقیق و راه‌حل‌های کمتر از حد بهینه در کاربردهای عملی می‌شوند. علاوه بر این، ماهیت تکرارشونده ذاتی در بیشتر روش‌های مدل‌محور، به ویژه برای بهینه‌سازی شبکه‌های بزرگ، باعث پیچیدگی محاسباتی و زمانی بالا می‌شود. این امر موجب می‌شود که زمان پردازش برخط به طور قابل ملاحظه‌ای طولانی شده و این روش‌ها نتوانند الزامات سخت‌گیرانه تأخیر پایین خدمات حساس به زمان در شبکه‌های نسل ششم را برآورده کنند. از این رو، اهمیت حیاتی وجود دارد که از ادغام اصول بهینه‌سازی و دانش دامنه با روش‌های یادگیری ژرف استفاده شود تا بر پیچیدگی‌های شبکه‌های نسل جدید غلبه گردد.

روش‌هایی نظیر یادگیری ژرف دانش‌محور و به ویژه گسترش ژرف به عنوان راهکارهای اصلی برای طراحی شبکه‌های بی‌سیم

<sup>14</sup> Loss Function

<sup>15</sup> Hyperparameter

<sup>16</sup> Weighted Minimum Mean Square Error

<sup>17</sup> Graph Neural Network

<sup>18</sup> Precoder

<sup>19</sup> Multiple Input Multiple Output

<sup>20</sup> Millimeter Wave

<sup>21</sup> Intelligent Reflecting Surface

<sup>22</sup> Algorithm Unfolding

<sup>23</sup> Energy Efficiency

<sup>24</sup> Nonconvex

<sup>25</sup> Sum-of-Ratios Problem

هوشمند مطرح شده‌اند. گسترش ژرف یک روش نمونه‌ای مدل-محور است که با تبدیل الگوریتم‌های تکرارشونده مرسوم، که از راه‌حل‌های نظری بهینه‌سازی شبکه مشتق شده‌اند به ساختارهای لایه‌ای از شبکه عصبی ژرف، تفسیرپذیری را بهبود می‌بخشد. این رویکرد امکان طراحی معماری‌های شبکه‌ای خاص مسئله و توابع فعال‌سازی سفارشی را فراهم می‌کند که برای مدیریت محدودیت‌های پیچیده سامانه‌های بی‌سیم بهینه شده‌اند. شبکه‌های مدل-محور حاصل از این روش، در مقایسه با معماری‌های متعارف یادگیری ژرف پارامترهای قابل آموزش کمتری دارند که منجر به کاهش زمان آموزش و پیچیدگی محاسباتی می‌شود. همچنین، این روش‌ها پیچیدگی و زمان اجرای مورد نیاز برای مسائل پردازش سیگنال، مانند تخصیص توان، را کاهش می‌دهند. بنابراین، برای توسعه سامانه‌های ارتباطی بی‌سیم کارآمدتر، تفسیرپذیرتر و انطباق‌پذیرتر در محیط‌های پویای نسل ششم، نیاز ضروری به روش‌هایی وجود دارد که دانش تخصصی دامنه را به‌طور کامل در معماری‌های یادگیری ژرف ادغام کنند.

## ۲.۱ چالش‌ها

محدودیت‌های کلی یادگیری ژرف در شبکه‌های بی‌سیم استفاده مستقیم از یادگیری ژرف صرفاً مبتنی بر داده را برای بهینه‌سازی شبکه‌های بی‌سیم نسل ششم دشوار می‌کند. وظایف مدیریت منابع و پردازش سیگنال در شبکه‌های بی‌سیم معمولاً با محدودیت‌های پیچیده و غیرمحدب متعددی مانند توابع درجه دوم، لگاریتمی یا کسری روبرو هستند که برآورده کردن دقیق آن‌ها یک چالش مهم و مداوم است.

کارایی شبکه‌های عصبی به شدت وابسته به حجم زیادی از داده‌های با کیفیت بالا و برچسب‌گذاری شده است، که جمع‌آوری آن‌ها در محیط‌های واقعی بی‌سیم به دلیل ملاحظات حریم خصوصی یا هزینه، اغلب غیرممکن یا زمان‌بر است. علاوه بر این، ماهیت جعبه سیاهی یادگیری ژرف و قابلیت تفسیرپذیری ضعیف آن، مانع از پذیرش گسترده آن در شبکه‌های ارتباطی می‌شود. مدل‌های یادگیری ژرف ذاتاً پرمصرف هستند و استقرار آن‌ها بر روی دستگاه‌های موبایل با منابع محاسباتی و انرژی محدود مانند عمر باتری چالش برانگیز است. چالش‌های مدل‌محور و مدل‌های گسترش ژرف روش‌های مدل‌محور مرسوم اغلب به دلیل مفروضات ساده‌سازی شده، منجر به مدل‌سازی غیردقیق می‌شوند. علاوه بر این، ذات تکرارشونده این روش‌ها پیچیدگی محاسباتی بالایی را برای بهینه‌سازی‌های بزرگ در نسل ششم ایجاد می‌کند. در مورد مدل‌های گسترش ژرف، توسعه آن‌ها با چالش‌های قابل توجهی روبروست، زیرا این روش یک طراحی مستقل نیست و باید بر اساس چارچوب‌های بهینه‌سازی تکرارشونده موجود مانند PGD<sup>26</sup> یا ISTA<sup>27</sup> ساخته شود. روش‌های بهینه‌سازی مرسوم ممکن است به حلقه‌های تودرتوی چند لایه نیاز داشته باشند که این با ساختار یکپارچه گسترش ژرف در تناقض است. همچنین، باز کردن عملیات‌های پیچیده در الگوریتم‌های تکرارشونده، مانند معکوس کردن ماتریس، در مدل‌های گسترش ژرف پیچیدگی بالایی به همراه دارد.

مدیریت محدودیت‌های پیچیده طراحی، مانند الزامات کیفیت خدمات (QoS<sup>28</sup>)، در مدل‌های گسترش ژرف دشوار است. مورد دیگر مسائل مربوط به هوش مصنوعی مولد<sup>29</sup> و مدل‌های زبانی بزرگ<sup>30</sup> است که در خصوص دومی یک مشکل عمده وجود دارد که آن تمایل مدل‌های زبانی بزرگ به توهم‌زایی<sup>31</sup> و تولید خروجی‌های فاقد دقت واقعی است، که می‌تواند قوانین فیزیکی و محدودیت‌های دنیای واقعی شبکه‌های بی‌سیم را نقض کند. همچنین، اکثر مدل‌های زبانی بزرگ برای پذیرش داده‌های متنی ساده طراحی

<sup>26</sup>Projected Gradient Decent

<sup>27</sup>Iterative Shrinkage-Thresholding Algorithm

<sup>28</sup>Quality of Service

<sup>29</sup>Generative Artificial Intelligence

<sup>30</sup>Large Language Model

<sup>31</sup>Hallucination



شده‌اند، که پردازش مستقیم انواع داده‌های متداول شبکه‌سازی مانند ترافیک سری زمانی یا شکل ساختاری گراف توسط آن‌ها را مختل می‌کند. استقرار مدل‌های زبانی بزرگ به طور کلی سربار محاسباتی و ذخیره‌سازی قابل توجهی به همراه دارد. در دسترسی چندگانه غیرمتعامد (<sup>32</sup>NOMA) اگر در مراحل اولیه لغو تداخل متوالی<sup>33</sup>، خطایی در رمزگشایی سیگنال‌های کاربر ضعیف رخ دهد، این خطا منجر به کاهش عملکرد رمزگشایی در مراحل بعدی می‌شود که به این پدیده انتشار خطا می‌گویند. همچنین، با افزایش تعداد کاربران و آنتن‌ها در شبکه‌های نسل بعدی، دستیابی به اطلاعات وضعیت کانال<sup>34</sup> به یک چالش بزرگ تبدیل می‌شود. امنیت لایه فیزیکی نیز یک چالش مهم است، به‌ویژه در مورد استراق سمع غیرفعال که در آن به دست آوردن اطلاعات وضعیت کانال برای اقدامات متقابل دشوار است.

## ۳.۱ ساختار گزارش

در این گزارش با تکیه بر تعدادی از پژوهش‌های جامع و پژوهش‌های نو مفاهیم و روش‌های پیشرفته یادگیری ژرف دانش‌محور و به‌طور خاص گسترش ژرف را در بهینه‌سازی شبکه‌های بی‌سیم به صورت نظام‌مند مورد ارزیابی قرار گرفته است. این گزارش شامل چهار بخش اصلی است که پس از این مقدمه (فصل ۱)، به مفاهیم پایه (فصل ۲) شامل مبانی یادگیری ژرف تعریف دقیق دانش دامنه و اصول گسترش ژرف، مرور کارهای پیشین (فصل ۳) براساس طبقه‌بندی کاربردهای لایه فیزیکی و ادغام دانش و در نهایت نتیجه‌گیری و کارهای آتی (فصل ۴) خواهد پرداخت. هدف این گزارش، ارائه اصول بنیادین و دستورالعمل‌هایی برای طراحی شبکه‌های بی‌سیم دانش‌محور است که به طور کامل از دانش دامنه خاص و روش‌های پیشرفته یادگیری ژرف بهره می‌برند، که نهایتاً منجر به توسعه سامانه‌های ارتباطی بی‌سیم کارآمدتر، قابل اعتمادتر و تفسیرپذیرتر می‌شود.

<sup>32</sup>Non Orthogonal Multiple Access

<sup>33</sup>Self-Interference Cancellation

<sup>34</sup>Channel State Information

## فصل ۲

# مفاهیم پایه‌ای

در این بخش با ارائه تعریفی روشن از دانش در حوزه شبکه‌های بی‌سیم شروع می‌کنیم. متعاقباً، این دانش حوزه خاص ارتباطات را به دو دسته کلی، یعنی دانش علمی و دانش تخصصی، طبقه‌بندی می‌کنیم. در هر دسته، انواع دانش خاص‌تر و همچنین بازنمایی‌ها و نمونه‌های معمول مربوط به آن‌ها، به تفصیل ارائه شده است.

درک مفاهیم پایه در حوزه شبکه‌های بی‌سیم هوشمند، به‌ویژه در نسل‌های آتی مانند نسل ششم، نیازمند فهم عمیق در خصوص چیستی دانش دامنه، ساختارهای نوین یادگیری ژرف و چگونگی ادغام این دو نظام فکری از طریق روش‌های مدل‌محور مانند گسترش ژرف است. این بخش به تشریح این اصول می‌پردازد. نخست، مفهوم و طبقه‌بندی دانش در شبکه‌های بی‌سیم است. دانش به معنای شناخت و خلاصه‌ای از نتایج اکتشافات بشری درباره جهان فیزیکی و ذهنی است که از دنیای واقعی کشف شده و سپس به عنوان اصول راهنما برای بهبود جهان به کار می‌رود. تعریف دقیق و یکنواخت دانش آسان نیست، زیرا دارای ماهیت انتزاعی و تعمیم‌یافته است و حتی در حوزه معرفت‌شناسی فلسفی همچنان موضوعی بحث‌برانگیز است. تعریف مرسوم ارائه شده توسط افلاطون بیان می‌کند که برای واجد شرایط بودن به عنوان دانش، یک گزاره باید سه معیار توجیه‌پذیری، حقیقی بودن و باورپذیری را برآورده کند. یک مفهوم جهانی‌تر از دانش، انباشت شناخت و تجربیات فردی است که از طریق فعالیت‌های عملی با هدف تغییر جهان عینی جمع‌آوری می‌شود و شامل درک ماهیت، ویژگی‌ها و حالات اشیا و همچنین روش‌های حل مسئله است. در هوش مصنوعی، دانش اغلب به توصیف روابط بین موجودیت‌ها در زمینه‌های خاص می‌پردازد و از طریق تجزیه و تحلیل مجموعه‌ای از اطلاعات به

دست می‌آید.

برای شبکه‌های بی‌سیم، دانش دامنه خاص ارتباطات<sup>۱</sup> به عنوان نظریه‌های فنی و شناخت تجربی تعریف می‌شود که توسط دانشمندان و متخصصان در فرآیند طراحی، ساخت و بهینه‌سازی شبکه‌های ارتباطی بی‌سیم انباشته شده است. این دانش دو جنبه محوری دارد. اول، در طول فرآیند مدل‌سازی، شامل توصیف شناختی ویژگی‌ها و قوانین تکاملی محیط شبکه و الزامات خدمت است همچون مدل‌سازی وظایف بی‌سیم مانند پردازش تصمیم‌گیری مارکوف<sup>۲</sup>. دوم، در طول فرآیند تصمیم‌گیری، خلاصه‌ای شناختی از اصول، الگوریتم‌ها و نظریه‌های مدل‌محور است که در حل مسئله به کار می‌روند مانند الگوریتم‌های نظری برای تخصیص توان.

دانش دامنه به دو دسته کلی دانش علمی و دانش تخصصی طبقه‌بندی می‌شود. این تمایز برای تفکیک پایه‌های بنیادی، اشکال نمایش و نقش‌های عملکردی آن‌ها در یادگیری ژرف دانش‌محور حیاتی است. دانش علمی با هدف ایجاد قوانینی جهانی است که سازوکار عملکرد شبکه‌های بی‌سیم را به صورت نظری توضیح می‌دهد. این دانش، نظام‌مند و عینی است و معمولاً به صورت صریح در قالب فرمول‌های ریاضی رسمی، الگوریتم‌ها و روابط احتمالی بیان می‌شود. در مقابل، دانش تخصصی به تجربیات عملی یا شهود اشاره دارد که توسط متخصصان و مهندسان در طول زمان خلاصه و اعتبارسنجی شده است. این دانش متنی، ذهنی و اغلب در عبارات نسبتاً غیررسمی مانند نمودارهای دانش یا گزاره‌های روایی ارائه می‌شود.

دسته‌بندی دیگر شبکه عصبی ژرف و نظام فکری یادگیری ژرف دانش‌محور است. یادگیری ژرف که یک فناوری حیاتی در دهه اخیر است، بر پایه شبکه عصبی ژرف بنا شده است. مدل‌های یادگیری ژرف شامل معماری‌های مختلفی هستند که باید بر اساس ویژگی‌های مسئله مورد نظر انتخاب شوند. برای مثال، دانش مربوط به همبستگی‌های زمانی ترافیک بی‌سیم یا وابستگی زمانی ناشی از تداخل بین نمادها<sup>۳</sup> در آشکارسازی سیگنال، استفاده از شبکه عصبی بازگشتی<sup>۴</sup> یا حافظه طولانی مدت-کوتاه مدت<sup>۵</sup> را هدایت می‌کند. همچنین، دانش درباره وابستگی فضایی و ساختار گرافی توپولوژی شبکه بی‌سیم، انتخاب شبکه عصبی گرافی را برای مسائلی مانند تخصیص توان توجیه می‌کند. یادگیری ژرف دانش‌محور روشی است که در آن دانش دامنه خاص ارتباطات به صورت صریح ادغام می‌شود تا داده‌های آموزشی ناکافی را تقویت کرده و طراحی مدل شبکه عصبی و الگوریتم‌های یادگیری را هدایت کند. این دانش یکپارچه‌شده، مستقل از داده‌های آموزشی اصلی برچسب‌گذاری شده است. یادگیری ژرف دانش‌محور به نظام‌فکری قابلیت تفسیر پیشینی تعلق دارد، به این معنی که قابلیت تفسیر از همان ابتدای طراحی یا فرآیند یادگیری مدل تعبیه می‌شود. این ویژگی در شبکه‌های بی‌سیم با الزامات ایمنی-بحرانی، بسیار حیاتی است. مسئله اساسی در یادگیری ژرف دانش‌محور چگونگی تعبیه مؤثر دانش در شبکه‌های عصبی است. یک طبقه‌بندی ساختاریافته از رویکردهای یکپارچه‌سازی دانش، بر اساس این که دانش به کدام یک از اجزای خط لوله یادگیری ژرف وارد می‌شود، ارائه شده است: این رویکردها شامل انتخاب مدل شبکه عصبی، سفارشی‌سازی مدل شبکه عصبی، ساختار تلفیق دانش و داده، طراحی تابع زیان، و پیکربندی فرآیند یادگیری است.

بخش آخر، سفارشی‌سازی مدل و گسترش ژرف است. سفارشی‌سازی مدل شبکه عصبی از دانش دامنه برای تنظیم مدل‌های شبکه عصبی ژرف با هدف افزایش تفسیرپذیری و انطباق‌پذیری استفاده می‌کند. این رویکرد به کاهش تعداد داده‌های آموزشی مورد نیاز و کاهش بار محاسباتی در آموزش آفلاین کمک می‌کند. این سفارشی‌سازی به سه دسته تقسیم می‌شود. طراحی زیرساختار، طراحی کل ساختار و شبکه‌های عصبی ترکیبی ساختار-محور. در طراحی زیرساختار<sup>۶</sup>، دانش در اجزای محلی شبکه مانند لایه

<sup>1</sup>Communication-specific domain knowledge

<sup>2</sup>Markov Decision Process

<sup>3</sup>Inter Symbol Interference

<sup>4</sup>Recurrent Neural Network

<sup>5</sup>Long Short-Term Memory

<sup>6</sup>Substructure Design

خروجی یا تابع فعال‌سازی گنجانده می‌شود تا تضمین کند خروجی‌ها به محدودیت‌های عملیاتی خاص مسئله پایبند هستند. به‌عنوان مثال، در مسأله پرتوسازی<sup>۷</sup> مقید MISO، از یک تابع افکنش<sup>۸</sup> خاص در لایه خروجی استفاده می‌شود تا مطمئن شویم خروجی در ناحیه امکان‌پذیر قرار می‌گیرد. طراحی کل ساختار یا همان گسترش الگوریتم، یک روش قدرتمند و مدل-محور است که دانش مربوط به الگوریتم‌های تکرارشونده اثبات شده را در کل ساختار شبکه عصبی تعبیه می‌کند. این روش، یک الگوریتم تکرارشونده مرسوم که از راه‌حل‌های نظری بهینه‌سازی شبکه مشتق شده است، را برای تعداد ثابتی از تکرارها باز می‌کند و هر تکرار را به عنوان یک لایه در شبکه عصبی ژرف در نظر می‌گیرد. گسترش ژرف، ساختار و منطق الگوریتم اصلی بهینه‌سازی را حفظ می‌کند و در عین حال پارامترهای ثابت الگوریتم را به پارامترهای انعطاف‌پذیر و قابل یادگیری در شبکه عصبی تبدیل می‌نماید. این رویکرد ساختار لایه‌ای قابل تفسیری را فراهم می‌کند که در آن هر لایه تقریباً متناظر با یک مرحله از الگوریتم تکرارشونده مدل-محور است. از آنجا که تعداد لایه‌ها یعنی تعداد تکرارها در مدل‌های گسترش ژرف ثابت است، این مدل‌ها دارای پیچیدگی محاسباتی ثابت و از پیش تعریف‌شده‌ای هستند. این ویژگی به زمان استنتاج<sup>۹</sup> سریع و قابل اعتماد به ویژه در کاربردهای آنی<sup>۱۰</sup> می‌انجامد. با ادغام صریح دانش دامنه در معماری، تفسیرپذیری شبکه بهبود می‌یابد. همچنین، این روش‌ها پارامترهای قابل آموزش کمتری نسبت به شبکه‌های جعبه سیاه عمومی دارند، که در نتیجه به مجموعه داده‌های آموزشی کمتری نیاز دارند. گسترش ژرف به طور گسترده برای حل مسائل پردازش سیگنال مانند آشکارسازی سیگنال، تخمین کانال و طراحی پرتوسازی و مدیریت منابع مانند تخصیص توان با استفاده از الگوریتم‌هایی نظیر WMMSE یا ADMM<sup>۱۱</sup>/PGD به کار رفته است. ساختار تلفیق دانش و داده و طراحی تابع زیان ساختار تلفیق دانش و داده یک نظام‌فکری ترکیبی است که در آن روش‌های نظری مدل-محور و شبکه‌های عصبی داده-محور به‌طور صریح و همکارانه برای حل مسائل بهینه‌سازی استفاده می‌شوند.

این ساختارها عمدتاً دانش علمی یا راه‌حل‌های نظری را به‌عنوان منبع اصلی دانش خود به کار می‌گیرند. این تلفیق می‌تواند به صورت ترتیبی یا موازی انجام شود. در حالت ترتیبی، بخش‌های دانش و داده به صورت متوالی عمل می‌کنند. این حالت خود به دو ساختار فرعی تقسیم می‌شود. نخست ساختار مختلط<sup>۱۲</sup> بوده که در آن مسئله اصلی به چندین زیرمسئله مرتبط تقسیم می‌شود. بخش‌های دانش به زیرمسائلی که راه‌حل‌های تحلیلی بهینه یا الگوریتم‌های نظری با پیچیدگی پایین دارند، می‌پردازند و بخش‌های داده زیرمسائلی را که روابط نگاشتی با ابعاد بالا یا مدل‌سازی دشوار دارند، حل می‌کنند. دوم ساختار پالایش<sup>۱۳</sup> است که بخش دانش یک راه‌حل تقریبی اولیه بر اساس مدل‌های نظری تعمیم‌یافته ارائه می‌دهد و سپس بخش داده با استفاده از داده‌های دنیای واقعی، این راه‌حل اولیه را پالایش می‌کند. این استفاده از راه‌حل‌های اولیه به‌طور قابل توجهی نیاز به تعداد نمونه‌های داده آموزشی برای بخش‌های داده را کاهش می‌دهد. مثال‌های کاربردی آن شامل تخمین کانال که یک تخمین‌گر مرسوم LS<sup>۱۴</sup> یک راه‌حل اولیه برای پالایش توسط یک شبکه عصبی ژرف ارائه می‌دهد و آشکارسازی سیگنال است. طراحی تابع زیان دانش‌محور روشی است که در آن دانش دامنه به صورت عبارت‌های اضافی فاقد برچسب در تابع زیان تعبیه می‌شود. این عبارات معمولاً از شناخت روش‌شناسی مدل‌سازی وظایف بی‌سیم و درک جامع از این وظایف نشأت می‌گیرند. هدف از این رویکرد، ترکیب یادگیری از داده‌های تجربی و دانش نظری برای بهبود قابلیت تعمیم و انطباق مدل است.

تابع زیان مقید-محور<sup>۱۵</sup> نوعی از تابع زیان است که دانش مربوط به محدودیت‌های ذاتی مسائل بهینه‌سازی را در تابع زیان

<sup>۷</sup>Beamforming

<sup>۸</sup>Projection Function

<sup>۹</sup>Inference

<sup>۱۰</sup>Real-Time

<sup>۱۱</sup>Alternating Direction Method of Multipliers

<sup>۱۲</sup>Mixing Structure

<sup>۱۳</sup>Refinement Structure

<sup>۱۴</sup>Least Square

<sup>۱۵</sup>Constraint-Specific Loss Function

گنجانده و با اعمال جریمه بر نقض محدودیت‌ها، فرآیند یادگیری را هدایت می‌کند. این رویکرد برای مدیریت مسائل با محدودیت‌های پیچیده از طریق شبکه‌های عصبی قدرتمند، روشی مؤثر است. یک مثال کلیدی، یادگیری اولیه-دوگانه<sup>۱۶</sup> است که با الهام از روش‌های دوگانگی لاگرانژی<sup>۱۷</sup>، محدودیت‌های پیچیده را به عنوان عبارت‌های جریمه خطی اضافی در تابع زیان گنجانده و بدین ترتیب یادگیری را هدایت می‌کند. همچنین، برای محدودیت‌های جمع خطی ساده، می‌توان محدودیت‌ها را با استفاده از توابع فعال‌سازی در تابع زیان تبدیل و اعمال کرد. پیکربندی فرایپارامترها دانش‌محور نیز از دانش قبلی برای تنظیم پارامترهای اولیه، ضریب‌های وزن‌دهی، و نرخ‌های یادگیری استفاده می‌کند. این کار به تسریع فرآیند آموزش و دستیابی به عملکرد قوی‌تر کمک می‌کند. به عنوان مثال، در الگوریتم‌هایی که از طریق گسترش ژرف حل می‌شوند، پارامترهای اولیه شبکه‌های گسترش داده شده می‌توانند از راه‌حل‌های الگوریتم‌های تکرارشونده نظری که بخشی از دانش دامنه هستند به ارث برده شوند و این مقداردهی اولیه عملکرد مدل را به طور قابل توجهی بهبود می‌بخشد.

<sup>16</sup>Primal-Dual Learning<sup>17</sup>Lagrange Duality



## فصل ۳

# مروری بر کارهای پیشین

۱.۳ تخمین کانال

۲.۳ آشکارسازی کانال

۳.۳ طراحی پیش کدگذار

۴.۳ سنجش و ارتباطات

۵.۳ رمزگشایی در کدهای تصحیح خطا

۶.۳ تخصیص توان

گسترش ژرف یک روش مبتنی بر یادگیری ژرف مدل-محور<sup>۱</sup> است که با تبدیل الگوریتم‌های تکرارشونده پیچیده به لایه‌های ساختاریافته در شبکه‌های عصبی ژرف، تفسیر پذیری را افزایش می‌دهد و دانش دامنه را به طور یکپارچه با یادگیری ژرف ادغام می‌کند. رویکرد

<sup>1</sup>Model-Driven

ترکیبی، از نقاط قوت هر دو روش بهره می‌برد تا وظایف پیچیده پردازش سیگنال در سامانه‌های ارتباطی را ساده‌سازی کند. در مقایسه با معماری‌های متعارف یادگیری ژرف، روش‌های مدل-محور پارامترهای قابل آموزش کمتری دارند و می‌توانند سریع‌تر و آسان‌تر، حتی با مجموعه داده‌های کوچک‌تر، آموزش داده شوند.

ایده اصلی گسترش ژرف، بهینه‌سازی پارامترهای کلیدی الگوریتم‌های مرسوم مانند اندازه‌ی گام در روش‌های مرسوم بر گرادین یا راه‌حل‌های اولیه بر اساس آموزش داده‌ها است. هر تکرار از الگوریتم تکرارشونده به عنوان یک لایه پنهان در شبکه عصبی ژرف گنجانده می‌شود و لایه‌ها روی هم چیده می‌شوند تا جریان بهینه‌سازی الگوریتم را شبیه‌سازی کنند. پژوهش‌های پیشین اغلب فاقد پوشش مطالعات جدید مرتبط با فناوری‌های حیاتی نسل ششم هستند، مانند طیف امواج میلی‌متری عظیم<sup>۲</sup>، MIMO، سطح بازتابنده هوشمند، و دسترسی چندگانه نسل بعدی و غالباً به حوزه‌های کاربردی محدودی در لایه فیزیکی می‌پردازند. مقاله با هدف رفع نیاز، یک بررسی عمیق از روش‌های گسترش ژرف اعمال شده در لایه فیزیکی ارتباطات بی‌سیم فراهم می‌آورد.

ساختار مقاله جامع، به چهار بخش اصلی تقسیم می‌شود. بخش A به ساختار کلی و گسترش ژرف می‌پردازد و مبانی یادگیری ژرف و معماری‌هایی مانند MLP<sup>۳</sup>، CNN<sup>۴</sup>، RNN<sup>۵</sup>، DRL<sup>۶</sup>، GNN<sup>۷</sup> و GAN<sup>۸</sup> را شرح می‌دهد. بخش B به صورت گسترده کاربردهای گسترش ژرف را در وظایف پردازش سیگنال لایه فیزیکی بررسی می‌کند. وظایف شامل آشکارسازی سیگنال مانند معماری DetNet مبتنی بر PGD در Massive MIMO، تخمین کانال شامل روش‌هایی مبتنی بر ISTA و MM<sup>۹</sup> در Massive MIMO و mmWave<sup>۱۰</sup>، طراحی پیش‌کدگذار که اغلب شامل گسترش دادن الگوریتم‌های WMMSE و PGD/PGA<sup>۱۱</sup> برای بهینه‌سازی نرخ جمع وزنی است، ISAC<sup>۱۲</sup>، کدگذاری برای کدهای تصحیح خطا مانند LDPC<sup>۱۳</sup> و توربو کدها با استفاده از ADMM یا BP<sup>۱۴</sup> گسترش داده شده، تخصیص توان (با استفاده از GNN برای پارامترسازی الگوریتم WMMSE، و امنیت لایه فیزیکی، مانند مقابله با اختلال هوشمند در Massive MIMO است. بخش C به طور خاص روش‌های مشترک در لایه فیزیکی را پوشش می‌دهد، مانند آشکارسازی و تخمین کانال مشترک که با معماری ComNet نشان داده می‌شود، آشکارسازی فعالیت و تخمین کانال مشترک، و تخمین کانال و طراحی پیش‌کدگذار مشترک مانند JCEHB<sup>۱۵</sup>. در نهایت، بخش D به بحث درباره چالش‌های اصلی توسعه مدل‌های گسترش ژرف و جهت‌گیری‌های تحقیقاتی آینده می‌پردازد. محدودیت‌ها عمدتاً از واقعیت ناشی می‌شوند که گسترش ژرف باید بر مبنای چارچوب‌های بهینه‌سازی متداول موجود ساخته شود.

چالش‌های اصلی مطرح شده عبارتند از: تضاد میان روش‌های بهینه‌سازی متعارف که اغلب به راه‌حل‌های لایه‌ای، تکرارشونده و حلقه‌های تودرتو نیاز دارند و مدل‌های گسترش ژرف که تمام متغیرها را به صورت موازی حل می‌کنند؛ دشواری در گسترش الگوریتم‌های اصلی که شامل عملیات‌های پیچیده مانند معکوس یا تجزیه ماتریس هستند و همچنین لزوم تعیین تعداد لایه‌ها به صورت پیش‌فرض که ممکن است تضمین همگرایی را در سناریوهای مختلف فراهم نکند؛ و چالش مدیریت قیود پیچیده طراحی مانند قیود کیفیت خدمت<sup>۱۶</sup> یا توابع هدف پیچیده مانند توابع کسری یا مرتبه بالا که تحمیل آن‌ها در لایه‌های شبکه عصبی ژرف دشوار است. همچنین در نگاشت‌های پیچیده‌ای که رابطه ورودی-خروجی واضحی ندارند مانند ردیابی پرتو با کمک حسگرهای

<sup>۲</sup>Massive Millimeter Wave

<sup>۳</sup>Multilayer Perceptron

<sup>۴</sup>Convolutional Neural Network

<sup>۵</sup>Recurrent Neural Network

<sup>۶</sup>Deep Reinforcement Learning

<sup>۷</sup>Graph Neural Network

<sup>۸</sup>Generative Adversarial Network

<sup>۹</sup>Multimodel

<sup>۱۰</sup>Millimeter Wave

<sup>۱۱</sup>Parallel Genetic Algorithm

<sup>۱۲</sup>Integrated Sensing and Communication

<sup>۱۳</sup>Low-Density Parity-Check

<sup>۱۴</sup>Belief Propagation

<sup>۱۵</sup>Joint Channel Estimation and Hybrid Beamforming

<sup>۱۶</sup>Quality of Service



چندوجهی، همچنان به مدل‌های جعبه سیاه یادگیری ژرف نیاز است.

برای غلبه بر محدودیت‌ها، جهت‌گیری‌های آینده شامل همزیستی گسترش ژرف با الگوریتم‌های حل مرسوم و مدل‌های جعبه سیاه به عنوان مثال، برای کاهش پیچیدگی زیرمسئله‌ها یا پیش‌پردازش داده‌های ورودی، توسعه مدل‌ها برای یادگیری چند-هدفه<sup>۱۷</sup> و ایجاد معماری‌های گسترش ژرف توزیع شده برای پلتفرم‌های دارای منابع محدود مانند دستگاه لبه<sup>۱۸</sup> به منظور کاهش تأخیر می‌شود. بررسی نتیجه می‌گیرد که گسترش ژرف، محدودیت‌های الگوریتم‌های مرسوم و مدل‌های یادگیری ژرف مستقل را با موفقیت برطرف کرده و به دلیل عملکرد با تأخیر کم، دقت بالا و تفسیرپذیری، نقشی اساسی در طراحی شبکه‌های بی‌سیم نسل بعدی ایفا خواهد کرد [۱].

مقاله مروری «پردازش سیگنال و یادگیری برای دسترسی چندگانه نسل بعدی در نسل ششم» یک ارزیابی جامع از تلاش‌های پژوهشی انجام شده برای ترکیب پردازش سیگنال پیشرفته و یادگیری ژرف در توسعه فناوری‌های NGMA<sup>۱۹</sup> در راستای تحقق سامانه‌های بی‌سیم نسل ششم است. پژوهشگران در بخش مقدمه توضیح می‌دهند که سامانه‌های ارتباطی بی‌سیم تا به امروز عمدتاً بر تعامد منابع تکیه داشته‌اند تا طراحی و پیاده‌سازی دسترسی و انتقال داده‌ها را تسهیل کنند. با حال، الزامات جدید و چالش‌برانگیز نسل ششم، مانند نیاز به نرخ داده اوج یک ترابیت بر ثانیه برای کاربردهایی نظیر VR<sup>۲۱</sup>/AR<sup>۲۰</sup> و همچنین پشتیبانی از اتصال انبوه مانند ۱۰ کاربر در هر کیلومتر مربع، مستلزم اصول طراحی انعطاف‌پذیرتری است که از تعامد فراتر می‌رود. پیشرفت‌های اخیر در پردازش سیگنال و یادگیری، رویکردهای امیدبخشی برای حل مسائل پیچیده و غیرقابل حل در NGMA ارائه داده‌اند. تمرکز اصلی مقاله بر دو فناوری محوری NGMA یعنی MRA<sup>۲۲</sup> و NOMA است.

مقاله با بررسی جامع روش‌های مبتنی بر مدل و مبتنی بر یادگیری در دو حوزه ادامه می‌یابد. در زمینه MRA، که برای سناریوهای اتصال انبوه مانند ارتباطات ماشین-محور انبوه حیاتی است، مشکل به عنوان بازایی پراکنده در سنجش فشرده<sup>۲۳</sup> مدل‌سازی می‌شود، چرا که تنها تعداد کمی از دستگاه‌ها در میان تعداد زیادی دستگاه، فعال هستند. روش‌های مبتنی بر مدل برای MRA شامل الگوریتم‌هایی مانند AMP ISTA و SBL هستند. در مقابل، روش‌های مبتنی بر یادگیری برای MRA، به ویژه گسترش ژرف، با تبدیل الگوریتم‌های تکرارشونده به لایه‌های شبکه عصبی، توانسته‌اند زمان محاسبات را به طور چشمگیری کاهش داده و دقت آشکارسازی/تخمین را افزایش دهند، زیرا پتانسیل غلبه بر خطاهای مدل‌سازی، ساختاری و همگرایی را دارند که در روش‌های مرسوم مبتنی بر مدل وجود دارد. در حوزه NOMA، که برای بهبود ظرفیت و کارایی طیفی از کدگذاری درهم‌نهی در فرستنده و حذف تداخل متوالی (SIC<sup>۲۴</sup>) در گیرنده استفاده می‌کند، چالش اصلی نیاز به اطلاعات دقیق وضعیت کانال است. پژوهش، طرح‌های عملی NOMA مانند PD-NOMA<sup>۲۵</sup> و CD-NOMA<sup>۲۶</sup> را مورد بحث قرار می‌دهد. یک یافته مهم است که NOMA نامتقارن، که در آن عدم همگام‌سازی کامل نمادها وجود دارد، می‌تواند برخلاف تصور، با افزایش تنوع نمونه‌برداری، ناحیه ظرفیت کانال دسترسی چندگانه را بزرگ‌تر کند و عملکرد سامانه‌های NOMA را بهبود بخشد. برای مدیریت مسائل غیرمحدب در NOMA، مانند تخصیص منابع، جفت‌سازی کاربران و طراحی فرستنده/گیرنده، از روش‌های یادگیری شامل یادگیری ژرف و یادگیری تقویتی<sup>۲۷</sup> استفاده می‌شود.

<sup>17</sup> Multi-Objective Learning

<sup>18</sup> Edge Device

<sup>19</sup> Next Generation Multiple Access

<sup>20</sup> Augmented Reality

<sup>21</sup> Virtual Reality

<sup>22</sup> Multiple Random Access

<sup>23</sup> Compressed Sensing

<sup>24</sup> Successive Interference Cancellation

<sup>25</sup> Power Domain Non Orthogonal Multiple Access

<sup>26</sup> Code Domain Non Orthogonal Multiple Access

<sup>27</sup> Reinforcement Learning

در ادامه پژوهش تعامل NGMA با چندین فناوری کلیدی نسل ششم را بررسی می‌کند. در ارتباط میدان نزدیک<sup>۲۸</sup>، که در به دلیل استفاده از آرایه‌های بزرگ و فرکانس‌های بالا اهمیت می‌یابد، ویژگی کانونی‌سازی پرتو<sup>۲۹</sup> امکان‌پذیر می‌شود. ویژگی برای NOMA بسیار مفید است، زیرا می‌تواند خوشه‌بندی کاربران را در حوزه فاصله ایجاد کند و همچنین امکان کدگذاری SIC از دور به نزدیک را فراهم می‌سازد که معمولاً در NOMA میدان دور غیرممکن است. ISAC به عنوان یک چارچوب دسترسی چندگانه دیده می‌شود که در آن کاربر ارتباطی و کاربر حسگری (هدف) منابع را به صورت غیر متعامد به اشتراک می‌گذارند و NOMA به عنوان ابزاری برای مدیریت تداخل بین وظیفه‌ای در سامانه‌های ISAC نامتعامد مورد استفاده قرار می‌گیرد. سطوح فرستنده و بازتابنده همزمان (STARS<sup>۳۰</sup>) که قابلیت ارسال و بازتاب سیگنال را دارند، در NOMA آزادی عمل بیشتری فراهم می‌کنند و با ایجاد تفاوت‌های کانالی مورد نیاز، عملکرد NOMA در T&R NOMA<sup>۳۱</sup> را بهبود می‌بخشند. علاوه بر، ارتباطات معنایی، که بر انتقال هدف و معنا تمرکز دارد، می‌تواند توسط NOMA برای کاربران متعدد فعال شود و همچنین می‌تواند با کاهش منابع رادیویی مانند توان ارسالی مورد نیاز برای دستیابی به یک هدف مشخص، عملکرد NOMA را به ویژه در سناریوهای نرخ سیگنال به نویز (SNR<sup>۳۲</sup>) پایین بهبود بخشد.

در بخش پایانی، پژوهشگران چالش‌های مهم در استفاده از روش‌های مبتنی بر یادگیری برای NGMA را مطرح می‌کنند. چالش‌ها شامل نیاز به ایجاد مجموعه داده‌های آموزشی واقع‌بینانه و سازگار با داده‌های آزمایشی، ضرورت طراحی شبکه‌های عصبی تخصصی و سبک‌وزن برای دستگاه‌های کم‌هزینه و منابع محدود با استفاده از روش‌هایی مانند گسترش ژرف و هرس مدل، و حل مشکل مقیاس‌پذیری و تعمیم‌پذیری مدل‌ها در سناریوهای ارتباطی متغیر مانند تغییرات در SNR یا تعداد کاربران فعال است. به طور کلی، پژوهش بر نقش حیاتی پردازش سیگنال هوشمند و یادگیری در ارائه راه‌حل‌هایی برای NGMA در عصر نسل ششم تأکید می‌کند [۲].

در پژوهش «بهینه‌سازی شبکه‌های بی‌سیم با گسترش ژرف: مطالعه تطبیقی بر دو مکانیزم گسترش ژرف» که یک مطالعه تطبیقی را بر روی دو مکانیزم گسترش ژرف برای انجام کارآمد کنترل توان در شبکه‌های بی‌سیم نسل بعدی انجام می‌دهد. با توجه به که پژوهش‌ها و توجهات زیادی به سمت نسل ششم معطوف شده است، که انتظار می‌رود الزامات فنی مانند نرخ داده بالا ۲۰ گیگابیت بر ثانیه برای پیوند فروسو، کیفیت تجربه بهتر، تأخیر کمتر و کارایی انرژی بسیار بالاتر (تا ۱۰۰ برابر بهتر از نسل پنجم) را برآورده سازد، کنترل توان با کارایی انرژی بالا یک حوزه حیاتی است.

پژوهش مشکل کنترل توان را به صورت به حداکثر رساندن WMMSE در شبکه‌های چند-سلولی با تداخل مدل‌سازی می‌کند. مسئله، که یک مسئله جمع نسبت‌ها (SoRP) و غیرمحدب است، به دست آوردن راه‌حل بهینه سراسری آن در زمان چندجمله‌ای دشوار است. برای حل مشکل، پژوهشگران از تبدیل برنامه‌ریزی کسری چندبُعدی استفاده می‌کنند تا مسئله را به دنباله‌ای از مسائل محدب تبدیل کنند. تبدیل با استفاده از متغیرهای کمکی انجام می‌شود تا تابع نرخ از کل مصرف توان و سیگنال از تداخل جدا شود. بر اساس چارچوب برنامه‌ریزی کسری، دو راه‌حل تکراری ارائه شده است: یک راه‌حل عددی (الگوریتم ۱) و یک راه‌حل بسته (الگوریتم ۲). هر دوی الگوریتم‌های تکرارشونده، اگرچه مزیت کاوش در فضای راه‌حل را دارند و طبق قضیه ۱ به یک نقطه ایستا از مسئله اصلی همگرا می‌شوند، اما زمان بر هستند و برای کاربردهای بلادرنگ در شبکه‌های نسل بعدی مناسب نیستند. برای رفع مشکل زمان محاسبات بالا در راه‌حل‌های تکراری، پژوهشگران دو مدل مبتنی بر یادگیری ژرف را بر اساس روش گسترش ژرف طراحی

<sup>28</sup>Near Field Communication<sup>29</sup>Beamfocusing<sup>30</sup>Simultaneously Transmitting And Reflecting Surface<sup>31</sup>Transmitting and Reflecting Non-Orthogonal Multiple Access<sup>32</sup>Signal to Noise Ratio

می‌کنند.

گسترش ژرف، که یک سازوکار مدل-محور است، دانش دامنه ارتباطات بی‌سیم را با قابلیت یادگیری یادگیری ژرف ترکیب می‌کند تا عملکرد را در مقایسه با روش‌های مرسوم و روش‌های صرفاً مبتنی بر داده بهبود بخشد. مدل اول، موسوم به MASUM<sup>33</sup>، بر اساس راه‌حل عددی (الگوریتم ۱) طراحی شده است. از آنجایی که راه‌حل نهایی الگوریتم ۱ به صورت عددی حل می‌شود و فرم بسته ندارد، MASUM به عنوان یک مدل نیمه-گسترش داده شده طراحی شده است که دانش دامنه از طریق تبدیل تکرارهای الگوریتم به زیرشبکه‌هایی برای محاسبه متغیرهای کمکی را با پیشرفت‌های یادگیری ژرف مبتنی بر داده ترکیب می‌کند.

مدل از دو خط لوله استفاده می‌کند؛ خط لوله اصلی تکرارهای الگوریتم را شبیه‌سازی می‌کند، و خط لوله دوم شامل زیرشبکه توجه<sup>34</sup> متعدد است که برای جبران افت عملکرد ناشی از لایه‌های مبتنی بر داده به کار می‌روند. مدل دوم، موسوم به FUM<sup>35</sup>، بر اساس راه‌حل بسته (الگوریتم ۲) طراحی شده است و به عنوان یک مدل کاملاً گسترش داده شده، می‌تواند تکرارها را به طور کامل شبیه‌سازی کند. مزیت وجود راه‌حل بسته است که مدل FUM می‌تواند بدون نیاز به لایه‌هایی مانند پیچشی یا توجه که در MASUM استفاده می‌شوند، ساخته شود و از دانش دامنه به صورت کامل بهره می‌برد. نتایج شبیه‌سازی نشان می‌دهد که هر دو مدل MASUM و FUM دقت بالا در تخمین کنترل توان را با سرعت استنتاج قابل توجهی در مقایسه با الگوریتم‌های تکرارشونده (الگوریتم ۱ و ۲) ارائه می‌دهند و ظرفیت بالقوه زیادی برای کاربردهای بلادرنگ در شبکه‌های نسل بعدی دارند. مدل FUM با بهره‌برداری کامل از دانش دامنه و بسته، عملکرد بهتری (تا 99.32% دقت) نسبت به MASUM (تا 98.80% دقت) در سناریوهای معمول و همچنین در برابر داده‌های خارج از آموزش نشان می‌دهد. همچنین مطالعه‌ای در مورد تأثیر تعداد لایه‌ها نشان داد که لزومی ندارد تعداد لایه‌ها برابر با تعداد تکرارهای مورد نیاز الگوریتم مرسوم برای همگرایی باشد؛ به عنوان مثال، FUM با تنها پنج یا شش لایه به عملکرد عالی دست می‌یابد که به میزان قابل توجهی سریع‌تر از اجرای ۱۴ لایه است.

در نهایت، پژوهشگران نتیجه می‌گیرند که وجود راه‌حل بسته، طراحی یک شبکه عصبی مدل-محور با گسترش کامل را تسهیل می‌کند، اما در صورت دشواری در گسترش کامل، می‌توان از رویکرد نیمه-گسترش داده شده استفاده کرد که در آن از پیشرفت‌های مدل‌های مبتنی بر داده برای جبران افت عملکرد استفاده می‌شود. کار آینده شامل طراحی شبکه‌های عصبی مدل-محور بهبود یافته برای مسائل با اندازه بزرگ با استفاده از معماری‌هایی مانند Inception-Residual و ادغام ایده شبکه عصبی مایع<sup>36</sup> برای محیط‌های پویا است [۳].

پژوهش «بیشینه‌سازی نرخ جمع عصبی برای شبکه‌های بی‌سیم مبتنی بر هوش مصنوعی: چارچوب روش ضرایب افزایشی متناوب و گسترش الگوریتم» یک رویکرد جدید به نام بیشینه‌سازی نرخ جمع عصبی<sup>37</sup> را برای حل مسائل غیرمحدب در زمینه‌ی بیشینه‌سازی نرخ جمع با در نظر گرفتن قیود توان کل برای دسترسی چندگانه‌ی پیوند فروسو معرفی می‌کند. پژوهش بر اهمیت بیشینه‌سازی نرخ جمع در شبکه‌های بی‌سیم برای بهبود کارایی، کاهش تداخل، و ارتقا تجربه‌ی کاربر تأکید دارد، به خصوص با توجه به چالش‌های فزاینده در شبکه‌های نسل بعدی که نیازمند طراحی‌های مبتنی بر هوش مصنوعی هستند. در حالی که روش‌های بهینه‌سازی مرسوم مانند تقریب محدب متوالی (SCA<sup>38</sup>) یا حداقل خطای میانگین مربعات وزن‌دار یا از نظر محاسباتی بسیار فشرده هستند یا تنها راه‌حل‌های زیربهینه ارائه می‌دهند، پژوهشگران پیشنهاد می‌کنند که ترکیب یادگیری ماشین با روش‌های بهینه‌سازی

<sup>33</sup>Multi-Attention-based Semi-Unfolding Model

<sup>34</sup>Attention Sub-Network

<sup>35</sup>Full-Unfolding Based Model

<sup>36</sup>Liquid Neural Networks

<sup>37</sup>Neural Sum Rate Maximization

<sup>38</sup>Successive Convex Approximation

مرسوم می‌تواند کارایی و اثربخشی را افزایش دهد.

پژوهش به جای که مستقیماً الگوریتم WMMSE را که شامل عملیات‌های پیچیده‌ای مانند معکوس ماتریس است و اجرای آن به عنوان لایه‌های شبکه عصبی دشوار است، گسترش دهد، یک روش جایگزین را پیشنهاد می‌دهد. رویکرد پژوهشگران ابتدا شامل توسعه‌ی یک روش سریع کاهش-بیشینه‌سازی<sup>۳۹</sup> است که از چارچوب تابع تداخل استاندارد<sup>۴۰</sup> و روش ضرایب افزایشی متناوب استفاده می‌کند. روش MM، اگرچه ممکن است همیشه بهینگی سراسری را تضمین نکند، اما یک بنیان محکم برای ترکیب روش‌های خاص دامنه با رویکردهای یادگیری ماشین فراهم می‌کند. سپس، با استفاده از روش گسترش الگوریتم، پژوهشگران تکرارهای الگوریتم MM مبتنی بر ADMM (الگوریتم ۱) را به لایه‌های شبکه عصبی قابل آموزش نگاشت می‌کنند تا مدل یادگیری شهودی‌تر و آسان‌تری برای آموزش ایجاد شود. نگاشت، که ساختارهای الگوریتم‌های تکرارشونده را حفظ می‌کند، منجر به افزایش بصیرت، تفسیر پذیری و سهولت آموزش می‌شود. معماری پیشنهادی (الگوریتم ۲)، شامل K لایه‌ی بیرونی است که وزن‌ها را یاد می‌گیرند و T لایه‌ی درونی که تکرارهای ADMM را برای حل زیرمسئله‌های تقریبی محدب باز می‌کنند. در داخل هر تکرار ADMM، به‌روزرسانی توان (p) نیز در D لایه باز می‌شود. پژوهشگران با استفاده از شبکه‌های عصبی برای ایجاد لایه‌های گسترش داده شده، برخی پارامترهای عملیات اصلی الگوریتم تکرارشونده مانند عملیات جمع را قابل آموزش می‌کنند.

نتایج عددی حاصل از شبیه‌سازی‌ها برتری قابل توجهی را در زمینه‌ی کارایی، عملکرد و تفسیرپذیری روش بیشینه‌سازی نرخ جمع عصبی پیشنهادی نسبت به الگوریتم‌های بهینه‌سازی نظری مرسوم (مانند WMMSE و SCA) و همچنین الگوریتم‌های مدرن مبتنی بر یادگیری (مانند GNN و TD3) نشان می‌دهد. به‌عنوان مثال، الگوریتم عصبی پیشنهادی در مقایسه با الگوریتم‌های یادگیری مبتنی بر داده مانند TD3 T، زمان آموزش کمتری نیاز دارد و در مقایسه با GNN، دقت بالاتری را حفظ می‌کند که منجر به عملکرد کلی بهتر می‌شود. مزایا، به ویژه برای بهینه‌سازی با قيود منابع در شبکه‌های بی‌سیم مبتنی بر هوش مصنوعی آینده، بسیار سودمند هستند. به طور خلاصه، مقاله با موفقیت یک چارچوب یادگیری ژرف مدل-محور برای حل یک مسئله بهینه‌سازی غیرمحدب حیاتی در شبکه‌های بی‌سیم ارائه داده است که هم از ساختار ریاضیاتی بهره می‌برد و هم از داده‌ها یاد می‌گیرد تا راه‌حلی سریع، دقیق و قابل استقرار در لایه‌ی AI-Native را ممکن سازد [۴].

پژوهشی دیگر در حوزه ORAN<sup>۴۱</sup> یادگیری ژرف مبتنی بر بهینه‌سازی را با هدف حل چالش غیرمحدب تخصیص مشترک زیر حامل و توان<sup>۴۲</sup> در شبکه‌های ORAN ارائه می‌دهد. هدف اصلی پژوهش، کمینه‌سازی کل مصرف توان است در حالی که اطمینان حاصل شود که کاربران نیازمندی‌های حداقل نرخ داده‌ی انتقال خود را برآورده می‌کنند. مسئله به دلیل غیرمحدب بودن و وجود قيود جفت‌شدگی، حل بهینه‌ی سراسری آن را با استفاده از الگوریتم‌های بهینه‌سازی مرسوم مانند SCA که اغلب به بهینه‌های محلی گیر می‌کنند و دارای هزینه‌های محاسباتی بالا هستند، با چالش مواجه می‌سازد.

پژوهشگران با معرفی معماری ORAN، بر ویژگی‌های آن از جمله رابط‌های باز، عناصر غیرمتمرکز، مجازی‌سازی سخت‌افزار و نرم‌افزار و کنترل هوشمند تأکید می‌کنند. معماری به گونه‌ای طراحی شده است که یک رابط AI-native را برای شبکه‌های بی‌سیم، از جمله سامانه‌های ماهواره‌ای-زمینی نوظهور، فراهم کند و یادگیری ژرف را به بخشی جدایی‌ناپذیر از عملکرد آن تبدیل کند. امر امکان بهبود کارایی طیفی، کاهش مصرف توان، تخصیص منابع پویا و کاهش هزینه‌های عملیاتی را فراهم می‌آورد. رویکرد اصلی پژوهش، طراحی مدل OpenRANet است که با مهندسی معکوس مشکل غیرمحدب به دنباله‌ای از زیرمسئله‌های محدب قابل حل

<sup>۳۹</sup>Majorization-Minimization

<sup>۴۰</sup>Standard Interference Function

<sup>۴۱</sup>Open Radio Access Network

<sup>۴۲</sup>Joint Subcarrier and Power Allocation

آغاز می‌شود. تبدیل با استفاده از روش‌هایی نظیر جداسازی، تبدیل متغیرها و آرام‌سازی محدب انجام می‌شود. پژوهشگران از چارچوب تابع تداخل استاندارد و ویژگی لگاریتم محدب بودن آن استفاده می‌کنند تا راه‌حل‌های اولیه-دوگانه را برای زیرمسئله‌های محدب استخراج کنند. راه‌حل‌ها به صورت الگوریتم تکرارشونده (الگوریتم ۱) با پیچیدگی محاسباتی پایین ارائه می‌شوند و تضمین همگرایی به یک بهینه محلی را بر اساس خواص تابع تداخل استاندارد و نظریه <sup>43</sup>KKT فراهم می‌کنند.

نوآوری OpenRANet در ادغام زیرمسئله‌های محدب به عنوان یک لایه‌ی بهینه‌سازی محدب در معماری یادگیری ژرف است. مدل از چهار جزء اصلی تشکیل شده است: یک فیلتر پیچشی برای استخراج ویژگی‌های سامانه‌های بزرگ مقیاس (برای مقابله با نفرین ابعاد)، لایه‌های متراکم، یک لایه‌ی پرتاب برای اعمال صریح قیود نرخ انتقال و لایه‌ی بهینه‌سازی محدب نهایی. لایه‌ی پرتاب، که تضمین می‌کند خروجی مدل (نرخ‌های انتقال) قیود نرخ مورد نیاز را برآورده می‌کند، یک جنبه حیاتی است که عملی بودن راه‌حل را تضمین می‌کند. لایه‌ی بهینه‌سازی محدب نیز خود شامل الگوریتم تکرارشونده اولیه-دوگانه (الگوریتم ۱) است که راه‌حل‌های بهینه برای زیرمسئله‌ی محدب (مسئله ۱۷) را فراهم می‌کند.

ادغام، که OpenRANet را به یک بهینه‌ساز یادگرفته شده تبدیل می‌کند، امکان می‌دهد تا پارامترهای اصلی بهینه‌سازی از طریق آموزش تنظیم شوند. معماری، با ترکیب دانش دامنه و یادگیری از داده‌ها، در مقایسه با روش‌های صرفاً مبتنی بر داده که نیاز به پارامترهای و داده‌های آموزشی بسیار زیادی دارند و روش‌های صرفاً مبتنی بر بهینه‌سازی که زمان محاسباتی بالایی دارند، کارایی و دقت را به طور قابل توجهی بهبود می‌بخشد. نتایج شبیه‌سازی نشان می‌دهد که OpenRANet توانسته است به دقت بسیار بالایی در تقریب زدن بهینه‌های سراسری دست یابد، به طوری که نتایج آن بسیار نزدیک به راه‌حل‌های به دست آمده از روش پرهزینه <sup>44</sup>BnB است. همچنین، در مقایسه با مدل‌های یادگیری ژرف صرفاً مبتنی بر داده مانند شبکه عصبی ژرف، GNN و DBN، OpenRANet به زمان آموزش به طور قابل ملاحظه‌ای کمتری نیاز دارد. همچنین OpenRANet در مقایسه با الگوریتم‌های بهینه‌سازی تکراری مانند SCA تعداد تکرارهای بسیار کمتری برای رسیدن به راه‌حل‌های بهینه نیاز دارد. امر بر اثربخشی چارچوب مدل-محور تأکید می‌کند.

پژوهشگران همچنین اهمیت لایه‌ی پرتاب را نشان دادند؛ در صورت حذف لایه، نرخ انتقال خروجی ممکن است به زیر حداقل سطح مورد نیاز کاهش یابد. در نهایت، پژوهشگران نتیجه‌گیری می‌کنند که OpenRANet می‌تواند به عنوان یک بنیان برای استراتژی‌های بهینه‌سازی بی‌سیم مبتنی بر هوش مصنوعی در سناریوهای گسترده‌تر، از جمله سامانه‌های چند-سلولی و شبکه‌های ORAN با نیازمندی‌های پیچیده‌تر مصرف توان مانند انرژی مصرفی برای پردازش سیگنال، خنک‌سازی و پردازش شبکه‌های عصبی عمل کند. همچنین، برای حفظ عملکرد در محیط‌های پویا، روش‌هایی مانند یادگیری انتقالی <sup>45</sup> و یادگیری تقویتی <sup>46</sup> می‌توانند در آینده با OpenRANet یکپارچه شوند [۳].

پژوهش دیگر به بررسی دسته گسترده‌ای از مسائل بهینه‌سازی تخصیص منابع به صورت غیرمتمرکز در شبکه‌های بی‌سیم می‌پردازد. این مسائل را می‌توان به عنوان یک مسئله یادگیری آماری با ساختار اطلاعاتی محلی <sup>47</sup> فرموله کرد که در آن فرستنده‌ها تنها به دانش محیط رادیویی محلی خود دسترسی دارند. با توجه به تغییرات سریع کانال که ویژگی بارز ارتباطات بی‌سیم است، سیاست‌های غیرمتمرکز از ابتدای کنترل توان مورد نیاز بوده‌اند. پژوهشگران یک رویکرد مبتنی بر یادگیری و مقیاس‌پذیر برای حل این دسته از

<sup>43</sup>Karush-Kuhn-Tucker<sup>44</sup>Branch and Bound<sup>45</sup>Transfer Learning<sup>46</sup>Incremental Learning<sup>47</sup>Localized Information Structure

مسائل تخصیص منابع غیرمتمرکز توسعه می‌دهند که هدف آن بهینه‌سازی یک مطلوبیت سراسری شبکه<sup>۴۸</sup> با رعایت قیود سامانه در محیط‌های غیرمتمرکز واقعی است. این چارچوب به‌طور خاص چالش‌های ذاتی شبکه‌های غیرمتمرکز، مانند تأخیر در تبادل اطلاعات و عدم همگام‌سازی ساعت‌های کاری دستگاه‌ها، را در نظر می‌گیرد. این عوامل محیطی انگیزه اصلی برای توسعه یک رویکرد یادگیری است که بتواند تأخیرها را تحمل کند، بدون همگام‌سازی عمل کند و از اطلاعات فراتر از همسایگی فوری یک گره بهره ببرد.

روش پیشنهادی از شبکه‌های عصبی گرافی تجمعی (Agg-GNN<sup>۴۹</sup>) استفاده می‌کند. Agg-GNNs با استفاده از تبادلات متوالی اطلاعات بین گره‌های همسایه، به دستگاه‌ها اجازه می‌دهند تا به‌طور محلی، اطلاعات مطلوبیت سراسری شبکه را انباشته کنند. این شبکه، دنباله‌ای از اطلاعات وضعیت تجمیع‌شده‌ی گرافی با تأخیر و احتمالاً ناهمگام را که به صورت محلی در هر فرستنده از طریق همسایگان چند-هاپ به دست می‌آید، پردازش می‌کند. در این معماری، لایه‌های متعدد پردازش اطلاعات با تأخیر پس از تجمیع سیگنال‌ها، امکان جمع‌آوری اطلاعات فضایی و زمانی مرتبط از شبکه بی‌سیم سراسری را فراهم می‌آورد. این پژوهش یک ویژگی ساختاری مهم را در سیاست تخصیص منابع به دست آمده اثبات می‌کند: تعادل جایگشتی<sup>۵۰</sup>. این خاصیت برای سیاست‌های تصمیم‌گیری خودمختار در شبکه‌های بی‌سیم بسیار حیاتی است، زیرا شکل ساختاری زیرین ذاتاً پویا است و اجازه می‌دهد که سیاست آموخته‌شده به شبکه‌های جدید با اندازه‌ها و توپولوژی‌های مختلف منتقل شود بدون از دست دادن بهینگی. قضیه ۱ بیان می‌کند که اگر دو شبکه جایگشت‌هایی از یکدیگر باشند، دارای همان فیلترهای بهینه‌ی Agg-GNN هستند. این امر انتقال‌پذیری<sup>۵۱</sup> Agg-GNN آموزش دیده به شبکه‌های بزرگتر یا شبکه‌هایی با پیکربندی‌های جدید را تسهیل می‌کند.

برای آموزش پارامترهای فیلتر Agg-GNN (تنسور فیلتر،  $A$ ) که باید به دقت تنظیم شوند تا عملکرد بهینه شود و قیود سامانه رعایت شود، پژوهشگران از یک روش یادگیری اولیه-دوگانه بدون مدل و بدون نظارت<sup>۵۲</sup> استفاده می‌کنند. این روش برای عملکرد در یک محیط ناهمگام و بدون نیاز به دانش صریح از مدل سامانه طراحی شده است و می‌تواند مسائل تخصیص منابع مقید را حل کند. در این چارچوب، قیود سامانه، مانند قیود توان کل، با استفاده از متغیرهای دوگان  $\mu$  و  $\lambda$  مدل Agg-GNN می‌تواند انواع مختلفی از مسائل تخصیص منابع، مانند تخصیص توان پویا در کانال‌های AWGN<sup>۵۳</sup> یا تخصیص توان با در نظر گرفتن تقاضای کاربران، یا دسترسی تصادفی توزیع‌شده، را پوشش دهد.

در شبیه‌سازی‌های عددی، پژوهشگران عملکرد Agg-GNN را در یک مسئله مرسوم تخصیص توان غیرمتمرکز در کانال AWGN با تداخل، با هدف بیشینه‌سازی نرخ جمع، آزمایش کردند. نتایج نشان داد که Agg-GNN عملکردی برتر نسبت به استراتژی‌های خط مبنای غیرمتمرکز دیگر مانند WMMSE در پیاده‌سازی غیرمتمرکز، تخصیص توان برابر، و تخصیص توان کامل تصادفی ارائه می‌دهد و عملکرد آن تقریباً با روش GNN متمرکز (Sel-GNN) در شبکه‌های بزرگ‌تر مطابقت دارد. این برتری در یک محیط ناهمگام نیز مشاهده شد، اگرچه نوسان بیشتری در منحنی‌های همگرایی وجود داشت. همچنین، توانایی انتقال‌پذیری مدل به شبکه‌هایی با اندازه برابر و شبکه‌های بزرگ‌تر با چگالی یکسان به صورت عددی تأیید شد، که نشان‌دهنده کارایی بیشتر Agg-GNN در اندازه پارامتر نسبت به شبکه‌های عصبی مرسوم است. به‌طور خلاصه، این مقاله یک چارچوب قدرتمند مبتنی بر Agg-GNN ارائه می‌دهد که نه تنها برای تخصیص منابع غیرمتمرکز و ناهمگام مناسب است، بلکه با حفظ خاصیت تعادل جایگشتی، مشکل مقیاس‌پذیری و

<sup>۴۸</sup>Global Network Utility<sup>۴۹</sup>Aggregation Neural Network<sup>۵۰</sup>Permutation Equivariance<sup>۵۱</sup>Transferability<sup>۵۲</sup>Unsupervised<sup>۵۳</sup>Additive White Gaussian Noise



انتقال‌پذیری مدل‌های یادگیری را در شبکه‌های بی‌سیم پویا به طور مؤثر حل می‌کند [۷].

پژوهش مروری «پیشرفت‌های اخیر در کنترل هوشمند مبتنی بر داده برای ارتباطات بی‌سیم: یک مرور جامع» یک بررسی جامع بر روی پیشرفت‌های اخیر در روش‌های مبتنی بر داده اعمال شده در شبکه‌های ارتباطات بی‌سیم، به‌ویژه با تمرکز بر تحولات پنج سال اخیر و کاربرد آن‌ها در اهداف کنترلی مختلف در سامانه‌های سایبر-فیزیکی بی‌سیم ارائه می‌دهد. پژوهشگران در مقدمه اظهار می‌دارند که ظهور سامانه‌های ارتباطی بی‌سیم نسل بعدی نویدبخش عصری است که با ویژگی‌های نرخ داده‌ی بالا، تأخیر کم، اتصال انبوه و کارایی انرژی برتر مشخص می‌شود. روش‌های مرسوم مبتنی بر بهینه‌سازی برای تأمین خواسته‌های پیچیده‌ی این سامانه‌های نوظهور ناکافی تشخیص داده شده‌اند. با افزایش حجم داده‌ها، ادغام روش‌های مبتنی بر داده، از جمله یادگیری ماشین<sup>۵۴</sup>، یادگیری ژرف، یادگیری تقویتی، یادگیری برخط<sup>۵۵</sup> و سایر روش‌های آماری، برای فعال‌سازی سازوکارهای کنترلی انطباق‌پذیر و هوشمند در سامانه‌های ارتباطی بی‌سیم آینده ضروری شده است.

این پژوهش بر این نکته تأکید دارد که قدرت اصلی مدل‌های مبتنی بر داده در توانایی آن‌ها برای یادگیری مستمر از داده‌های بلادرنگ است، که به آن‌ها امکان می‌دهد به طور پویا با شرایط متغیر شبکه سازگار شوند. این قابلیت انطباق‌پذیری، مدیریت کارآمدتر منابع، مانند استفاده بهتر از طیف و بهینه‌سازی انرژی، را تضمین می‌کند. علاوه بر انطباق‌پذیری، مقیاس‌پذیری مزیت مهم دیگری است، زیرا رویکردهای مبتنی بر داده می‌توانند پیچیدگی شبکه‌های مدرن مانند MIMO Massive را با یادگیری الگوهای داده‌ای که روش‌های مرسوم قادر به مدیریت آن‌ها نیستند، مدیریت کنند.

پژوهش حاضر، بر خلاف پژوهش‌های موجود که اغلب محدود به موضوعات خاصی هستند یا صرفاً بر مدیریت منابع لایه اطلاعاتی تمرکز دارند، بحث را به حوزه‌های حیاتی مقابل در سامانه‌های سایبر-فیزیکی بی‌سیم گسترش می‌دهد؛ انطباق‌پذیری پیوند<sup>۵۶</sup> به‌صورت پویا به‌تنظیم پارامترهای پیوند، مانند شاخص MCS و سطح توان، برای بهینه‌سازی عملکرد تحت شرایط کانالی متغیر می‌پردازد. زمان‌بندی کاربر<sup>۵۷</sup> به تخصیص کارآمد منابع شبکه به کاربران برای بهینه‌سازی توان عملیاتی و کمینه‌سازی تأخیر می‌پردازد. تخصیص طیف<sup>۵۸</sup> به بهبود کارایی طیف و کاهش تداخل می‌پردازد. مدیریت پرتو<sup>۵۹</sup> به بهینه‌سازی شکل‌دهی و مدیریت پرتو برای افزایش کیفیت سیگنال و کارایی ارتباطات mmWave.

کنترل توان<sup>۶۰</sup> که به استفاده از رویکردهای مبتنی بر داده برای کنترل توان پویا به منظور بهینه‌سازی کارایی انرژی با حفظ کیفیت ارتباط می‌پردازد. طراحی مشترک سامانه ارتباطات و کنترل<sup>۶۱</sup> به ادغام فرآیندهای فیزیکی صنعتی با انتقال اطلاعات، که برای عملیات اینترنت اشیا حیاتی است می‌پردازد.

در حوزه‌ی انطباق‌پذیری پیوند، روش‌های مبتنی بر یادگیری می‌توانند مدل‌هایی را مستقیماً آموزش دهند تا عملکرد انتقال را بر اساس وضعیت کانال پیش‌بینی کنند یا پارامترهای روش‌های مرسوم مانند OLLA<sup>۶۲</sup> را بهینه سازند تا سریع‌تر از روش‌های مرسوم مانند MINP<sup>۶۳</sup>، با شرایط کانال متغیر سازگار شوند. در کنترل توان، روش‌هایی مانند شبکه‌های عصبی ژرف، شبکه‌های عصبی گرافی و شبکه‌های عصبی بازگشتی برای مدل‌سازی روابط غیرخطی پیچیده و دستیابی به کارایی انرژی و طیفی بهتر نسبت به روش‌های مرسوم WMMSE استفاده می‌شوند. به‌ویژه، GNN‌ها به دلیل توانایی‌شان در مدل‌سازی ساختار شکل‌ساختاری شبکه‌های بی‌سیم،

<sup>54</sup>Machine Learning

<sup>55</sup>Online Learning

<sup>56</sup>Link Adaptation

<sup>57</sup>User Scheduling

<sup>58</sup>Spectrum Allocation

<sup>59</sup>Beam Management

<sup>60</sup>Power Control

<sup>61</sup>Co-design of Communication and Control System

<sup>62</sup>Outer Loop Link Adaptation

<sup>63</sup>Mixed-Integer Nonlinear Programming

برای تخصیص منابع مناسب هستند. برای مدیریت پرتو، روش‌های مبتنی بر داده مانند یادگیری ژرف،<sup>64</sup> UCB و<sup>65</sup> TS به دلیل توانایی در یادگیری الگوهای پیچیده و مقابله با ماهیت پویا کانال‌های mmWave، برای وظایفی مانند انتخاب پرتو<sup>66</sup> و ردیابی پرتو<sup>67</sup> استفاده می‌شوند.

در نهایت، پژوهشگران چالش‌های کلیدی موجود در الگوریتم‌های مبتنی بر داده را مطرح می‌کنند. این چالش‌ها شامل مسائل مربوط به انطباق‌پذیری پویا در برابر تغییرات مستمر در نیازهای کاربران و محیط، امنیت مدل‌های هوش مصنوعی در برابر حملات خصمانه و مسموم‌سازی داده و سربار محاسباتی و سخت‌افزاری مدل‌های پیچیده یادگیری ژرف هستند. همچنین، برای تضمین موفقیت در آینده، نیاز به تحقیق در زمینه‌ی توسعه مدل‌های سبک‌وزن یادگیری ماشین، افزایش تعمیم‌پذیری الگوریتم‌ها و انتقال‌پذیری سیاست‌های آموخته‌شده به اشکال ساختاری شبکه‌ای جدید است. این پژوهش یک تحلیل انتقادی از این چالش‌ها و بینش‌هایی در مورد راه‌حل‌های بالقوه و جهت‌گیری‌های تحقیقاتی آینده ارائه می‌دهد، از جمله بحث در مورد یکپارچه‌سازی با نسل ششم و سیستم‌های حسگری-ارتباطی یکپارچه [۸].

پژوهش مروری «یک بررسی جامع از یادگیری ژرف برای بهینه‌سازی شبکه بی‌سیم هوشمند در نسل ششم» یک تحلیل جامع و ساختاریافته از نظام‌فکری یادگیری ژرف دانش‌محور ارائه می‌دهد که به طور خاص برای بهینه‌سازی شبکه‌های بی‌سیم هوشمند در عصر نسل ششم طراحی شده است. این شبکه طبق پیش‌بینی‌ها زیرساخت اساسی برای یک جهان رقمی کاملاً هوشمند، فراگیر و پایدار خواهد بود و هدف آن ادغام عمیق ارتباطات، محاسبات، سنجش و هوش بومی برای پشتیبانی از تعاملات بی‌وقفه در زمان واقعی بین انسان‌ها، ماشین‌ها و محیط‌ها است. نسل ششم با ویژگی‌هایی مانند پوشش وسیع، خدمات متنوع در تمام سناریوها، اتصال انبوه و ناهمگنی پویا مشخص می‌شود، که این ویژگی‌ها منجر به مسائل بهینه‌سازی بزرگ‌مقیاس و بسیار پیچیده در شبکه می‌شوند.

پژوهشگران ابتدا محدودیت‌های روش‌های مرسوم مدل-محور و روش‌های صرفاً داده-محور را برای توجیه نیاز به رویکرد دانش‌محور بررسی می‌کنند. روش‌های مدل-محور نظری بر مبنای دانش دامنه مانند نظریه ارتباطات و اصول ریاضیاتی طراحی می‌شوند و قابلیت تضمین عملکرد بالایی دارند و رفتار شبکه را قابل فهم می‌سازند. با این حال، در مسائل پیچیده نسل ششم، این روش‌ها اغلب به دلیل تکیه بر مفروضات ساده‌سازی شده مانند اختلال گاوسی و رفتار سامانه خطی و ماهیت تکراری ذاتی خود، از شدت محاسباتی بالا و زمان پردازش طولانی رنج می‌برند، که آن‌ها را برای تأخیر کم مورد نیاز نسل ششم نامناسب می‌سازد. از طرف دیگر، روش‌های صرفاً داده-محور یادگیری ژرف، به دلیل قابلیت تقریب سراسری خود، می‌توانند روابط پیچیده را مدیریت کرده و پس از آموزش برون‌خط، سرعت استنتاج بالایی را برای کاربردهای بلندرنج ارائه دهند. اما، این مدل‌ها با چالش‌هایی مانند کمبود داده‌های برچسب‌دار با کیفیت بالا، دشواری در مدیریت قیود چندگانه در تخصیص منابع، و عمل کردن به عنوان «جعبه سیاه» به دلیل فقدان تفسیرپذیری و تضمین عملکرد، مواجه هستند که کاربرد آن‌ها را در سناریوهای حیاتی مانند رانندگی خودران متصل محدود می‌کند. یادگیری ژرف دانش‌محور برای رفع این محدودیت‌ها توسعه یافته است. این رویکرد با ادغام صریح دانش دامنه، که بنیان روش‌های مدل-محور است، در شبکه‌های عصبی داده-محور، از نقاط قوت هر دو روش بهره می‌برد. این کار به طور قابل توجهی تفسیرپذیری مدل را بهبود می‌بخشد، تعداد پارامترهای قابل یادگیری و نمونه‌های داده‌ی آموزشی مورد نیاز را کاهش می‌دهد، و همچنان سرعت استنتاج سریع شبکه‌های عصبی را حفظ می‌کند، که برای الزامات تأخیر کم در شبکه‌های پویا ضروری است.

<sup>64</sup>Upper Confidence Bound<sup>65</sup>Thompson Sampling<sup>66</sup>Beam Selection<sup>67</sup>Beam Tracking



تفسیرپذیری در این رویکرد از نوع قابلیت تفسیر پیشینی است که تفسیرپذیری را از همان ابتدای فرآیند طراحی مدل یا یادگیری جاسازی می‌کند. این پژوهش با مشاهده اینکه پژوهش‌های موجود فاقد یک تعریف صریح و طبقه‌بندی نظام‌مند برای رویکردهای ادغام دانش در بهینه‌سازی شبکه‌های بی‌سیم هستند، سه سهم اصلی را برای پر کردن این شکاف‌ها ارائه می‌دهد.

سهم اول، ارائه یک تعریف خلاقانه از دانش حوزه خاص ارتباطات است. این دانش شامل نظریه‌های فنی و شناخت تجربی انباشته شده توسط دانشمندان و متخصصان در فرآیند طراحی، ساخت و بهینه‌سازی شبکه‌های بی‌سیم است. این دانش به طور گسترده به دو دسته تقسیم می‌شود. نخست دانش علمی، که شامل قوانین انتقال نظری مانند سه قضیه شانون، روش‌های مدل‌سازی و راه‌حل‌های نظری مانند پیش‌کدگذاری<sup>68</sup> ZF یا الگوریتم‌های تکرارشونده مانند<sup>69</sup> WF است و معمولاً به صورت صریح در فرمول‌های ریاضیاتی و الگوریتم‌های تکرارشونده بیان می‌شود. دسته دوم، دانش خبرگی<sup>۷۰</sup> است که شامل تجربه عملی یا شهود جمع‌آوری شده توسط متخصصان است و شامل پروتکل‌های شبکه مانند 3GPP و IEEE 802.11، روابط تجربی بین نهادها مانند گراف‌های دانش و ویژگی‌های منحصر به فرد شبکه‌های بی‌سیم مانند همبستگی زمانی یا شکل‌ساختاری گرافی فضایی می‌شود. سهم دوم، ارائه یک طبقه‌بندی جدید و پیشرو برای رویکردهای ادغام دانش در شبکه‌های بی‌سیم است. این طبقه‌بندی بر اساس این که دانش در کدام جزء از خط لوله یادگیری ژرف و چگونه ادغام می‌شود، ساختار یافته است. رویکرد اول انتخاب مدل شبکه عصبی دانش‌محور است که استفاده از دانش ویژگی‌های منحصر به فرد شبکه مانند همبستگی زمانی یا شکل‌ساختاری گرافی، برای انتخاب مدل مناسب مانند RNN برای همبستگی زمانی یا GNN برای شکل‌ساختاری گرافی. رویکرد دوم سفارشی‌سازی مدل شبکه عصبی دانش‌محور از طریق تغییر ساختار مدل با استفاده از دانش دامنه. این رویکرد به سه زیردسته تقسیم می‌شود؛ اول طراحی زیرساختار مانند افزودن لایه‌ی پرتاب به لایه‌ی خروجی برای رعایت قیود ریاضیاتی، دوم طراحی ساختار کامل مانند گسترش الگوریتم‌های تکرارشونده مانند PGD یا WMMSE به لایه‌های شبکه عصبی و دسته‌بندی سوم ساختاردهی شبکه‌های عصبی ترکیبی است که ساختار جریان الگوریتم‌های مدل-محور چندبloomکی مانند<sup>71</sup> AM را حفظ می‌کند. رویکرد بعدی ساختاردهی معماری ترکیب دانش و داده است که از تلفیق صریح بخش‌های مدل-محور یا بخش‌های دانش و بخش‌های داده-محور یا شبکه‌های عصبی در حالت ترکیبی مانند ساختار ترکیبی برای تقسیم مسائل به زیرمسئله‌های مختلف، یا ساختار پالایشی که راه‌حل اولیه‌ی مدل-محور توسط یادگیری ژرف پالایش می‌شود یا موازی برای افزایش پایداری و تاب‌آوری سامانه. رویکرد دیگر طراحی تابع زیان دانش‌محور که از افزودن عباراتی به تابع زیان که قیود یا ویژگی‌های خاص مسئله را جریمه می‌کند. این شامل تابع زیان خاص قیود مانند استفاده از تکنیک دوگان لاگرانژی برای جریمه کردن قیود پیچیده در یادگیری اولیه-دوگان و تابع زیان خاص ویژگی مانند ایجاد پاداش‌های فوری مبتنی بر دانش برای حل مسئله پراکندگی پاداش در یادگیری تقویتی عمیق است. رویکرد آخر، پیکربندی فرایارمتر دانش‌محور است که استفاده از دانش دامنه برای تنظیم پارامترهای مدل‌محور مانند طراحی ورودی و خروجی شبکه‌های عصبی برای کاهش بُعد برای دور زدن «نفرین ابعاد» در یادگیری تقویتی عمیق یا پارامترهای الگوریتم‌محور مانند مقداردهی اولیه پارامترهای قابل آموزش در شبکه‌های گسترش داده شده بر اساس مقادیر نظری الگوریتم. سهم سوم، بررسی جامع ادبیات مربوط به تخصیص منابع و پردازش سیگنال بر اساس این طبقه‌بندی جدید است.

در پایان، پژوهشگران به چندین چالش کلیدی پیش روی یادگیری ژرف دانش‌محور اشاره می‌کنند، از جمله نیاز به توسعه‌ی روش‌های یادگیری ژرف دانش‌محور برای مدیریت سختگیرانه‌ی قیود پیچیده‌ی غیرخطی مانند توابع کسری و لگاریتمی که در

<sup>68</sup>Zero-Force<sup>69</sup>Water-Filling<sup>70</sup>Expert Knowledge<sup>71</sup>Alternating Minimization

شبکه‌های بی‌سیم وجود دارند؛ بهبود تفسیرپذیری و تضمین‌های عملکردی مدل‌های دانش‌محور برخلاف روش‌های مدل-محور مرسوم؛ ظرفیت ادغام هوش مصنوعی مولد و مدل‌های زبان بزرگ در شبکه‌های بی‌سیم با وجود چالش‌هایی مانند تمایل مدل‌های زبانی به تولید خروجی‌های توهم‌زایی و عدم توانایی آن‌ها در پردازش مستقیم انواع داده‌های غیرمتنی شبکه‌های بی‌سیم و مسائل عملیاتی مانند سازگاری با شیوه‌نامه‌های موجود و محدودیت‌های سخت‌افزاری دستگاه‌های سیار که برای استقرار موفقیت‌آمیز در شبکه‌های واقعی ضروری هستند. این بررسی به عنوان یک راهنمای روشن‌گر برای استفاده مؤثر از دانش دامنه در طراحی شبکه‌های نسل ششم هوشمند، کارآمد و قابل اعتماد عمل می‌کند [۱۰].

پژوهش دیگری به یک طرح ابتکاری برای طراحی شکل موج دوگانه-عملکردی<sup>۷۲</sup> در سامانه‌های ISAC کمک‌شده با RIS<sup>۷۳</sup> ارائه می‌دهد که مبتنی بر روش یادگیری گسترش ژرف است. پژوهشگران در مقدمه اشاره می‌کنند که رشد تصاعدی ترافیک داده‌های بی‌سیم، ازدحام طیفی را به طور قابل توجهی تشدید کرده و ISAC به عنوان یک روش امیدوارکننده برای کاهش این مشکل از طریق اشتراک‌گذاری طیف و سخت‌افزار برای ارائه خدمات دوگانه ارتباطی و حسگری، در حال ظهور است. با این حال، طرح‌های ISAC، به‌ویژه آن‌هایی که شامل بهینه‌سازی مشترک هستند، معمولاً با چالش پیچیدگی محاسباتی بالا مواجه می‌شوند. این امر استقرار برخط<sup>۷۴</sup> را دشوار می‌سازد، که نیاز به راه‌حل‌های سریعی دارد.

برای حل این مشکل، پژوهشگران از قابلیت‌های قدرتمند برازش غیرخطی و سرعت استنتاج سریع یادگیری ژرف بهره می‌برند و یک طرح شکل موج دوگانه-عملکردی را بر اساس یادگیری گسترش ژرف پیشنهاد می‌کنند. هدف بهینه‌سازی، کمینه‌سازی مجموع وزن دار انرژی تداخل چندکاربره و اختلاف شکل موج از طریق طراحی مشترک شکل موج ارسالی و تغییر فاز در RIS است. این مسئله یک مشکل بهینه‌سازی غیرمحدب و چالش برانگیز است که متغیرهای تزویج‌شده‌ای دارد.

برای رسیدگی به این مسئله، پژوهشگران ابتدا یک الگوریتم تکرارشونده مبتنی بر روش ضرایب افزایشی متناوب توسعه می‌دهند تا مشکل غیرمحدب را به دنباله‌ای از چهار زیرمسئله تقسیم کرده و به طور متناوب آن‌ها را حل کند. راه‌حل بسته برای هر زیرمسئله استخراج می‌شود و این الگوریتم تضمین می‌کند که به یک نقطه همگرایی دست یابد. با این حال، پیچیدگی محاسباتی این الگوریتم تکرارشونده ADMM بالاست و به دلیل عملیات‌هایی مانند معکوس ماتریس در به‌روزرسانی شکل موج و تغییر فاز، همچنان مانع بزرگی برای استقرار برخط است. برای کاهش بیشتر پیچیدگی محاسباتی، پژوهشگران یک شبکه‌ی عصبی مدل-محور به نام ADMM-NET طراحی می‌کنند. ADMM-NET، الگوریتم تکرارشونده ADMM پیشنهادی را به یک معماری لایه‌به‌لایه باز می‌کند و عملیات‌های پیچیده‌ی معکوس ماتریس را با تقریب‌هایی با پیچیدگی پایین‌تر که شامل پارامترهای قابل یادگیری هستند، جایگزین می‌کند. گسترش ژرف با ادغام دانش دامنه در ساختار شبکه عصبی ژرف، نه تنها تفسیرپذیری شبکه را به طور قابل توجهی افزایش می‌دهد، بلکه مزایای دیگری مانند نیاز کمتر به داده‌های آموزشی و پارامترهای شبکه‌ی کمتر را نسبت به شبکه‌های جعبه سیاه ارائه می‌کند. ADMM-NET در یک روش یادگیری بدون نظارت آموزش داده می‌شود تا به جای محدود شدن توسط راه‌حل‌های زیربهینه‌ی تولید شده توسط الگوریتم ADMM، به جستجوی یک راه‌حل با کیفیت بالاتر بپردازد. تابع زیان اصلاح‌شده‌ای نیز برای تضمین غیرمنفی بودن و افزایش کارایی آموزش استفاده می‌شود.

نتایج شبیه‌سازی نشان می‌دهد که ADMM-NET در مقایسه با الگوریتم تکرارشونده ADMM، از نظر پیچیدگی محاسباتی و عملکرد برتری دارد، که استقرار برخط را تسهیل می‌کند. همچنین، در مقایسه با شبکه‌ی عصبی جعبه سیاه که به عنوان معیار

<sup>۷۲</sup>Dual-Functional Waveform<sup>۷۳</sup>Reconfigurable Intelligent Surface<sup>۷۴</sup>Online Deployment

استفاده شده است که هم در حالت نظارت شده و هم بدون نظارت آزمایش شده، ADMM-NET عملکرد بهتری را در نرخ خطای بیت و نرخ قابل دستیابی متوسط، به ویژه در مقایسه با یادگیری نظارتی که توسط عملکرد ADMM محدود شده است، نشان می‌دهد. ADMM-NET با ترکیب دانش دامنه و داده‌های آموزشی، می‌تواند بر نقاط ضعف شبکه‌های جعبه سیاه مانند گیر افتادن در بهینه‌های محلی در حالت بدون نظارت و نیاز به نمونه‌های آموزشی عظیم غلبه کند. به طور کلی، ADMM-NET بهترین تعادل را بین عملکرد تبادل MUI انرژی و اختلاف شکل موج در میان تمام الگوریتم‌های مورد بررسی ایجاد می‌کند [۱۳].

## فصل ۴

# نتیجه‌گیری و کارهای آینده

در این فصل نخست در **بخش ۱.۴**، به‌عنوان نتیجه‌گیری مروری بر آنچه که گفته شد، صورت خواهد پذیرفت. در ادامه نیز در **بخش ۲.۴**، مسائل باز این حوزه و پیشنهاداتی برای کارهای آینده، ارائه خواهد شد.

۱.۴ نتیجه‌گیری

۲.۴ کارهای آینده

- [1] S. Deka, K. Deka, N. T. Nguyen, S. Sharma, V. Bhatia, and N. Rajatheva, "Comprehensive review of deep unfolding techniques for next-generation wireless communication systems," 2025.
- [2] W. Chen, Y. Liu, H. Jafarkhani, Y. C. Eldar, P. Zhu, and K. B. Letaief, "Signal processing and learning for next generation multiple access in 6g," *IEEE Journal of Selected Topics in Signal Processing*, vol.18, no.7, pp.1146–1177, 2024.
- [3] A. Adam, M. A. Elhassan, and E. M. Diallo, "Optimizing wireless networks with deep unfolding: Comparative study on two deep unfolding mechanisms," *arXiv preprint arXiv:2403.18930*, 2024.
- [4] S. Chen and C. W. Tan, "Neural sum rate maximization for ai-native wireless networks: Alternating direction method of multipliers framework and algorithm unrolling," in *Proceedings of the 2nd International Workshop on Networked AI Systems*, NetAISys '24, (New York, NY, USA), p.13–18, Association for Computing Machinery, 2024.
- [5] S. Chen, C. W. Tan, X. Zhai, and H. V. Poor, "Openranet: Neuralized spectrum access by joint subcarrier and power allocation with optimization-based deep learning," 2025.
- [6] A. Chowdhury, G. Verma, C. Rao, A. Swami, and S. Segarra, "Unfolding wmmse using graph neural networks for efficient power allocation," *IEEE Transactions on Wireless Communications*, vol.20, no.9, pp.6004–6017, 2021.
- [7] Z. Wang, M. Eisen, and A. Ribeiro, "Learning decentralized wireless resource allocations with graph neural networks," *IEEE Transactions on Signal Processing*, vol.70, pp.1850–1863, 2022.
- [8] W. Huo, H. Yang, N. Yang, Z. Yang, J. Zhang, F. Nan, X. Chen, Y. Mao, S. Hu, P. Wang, X. Zheng, M. Zhao, and L. Shi, "Recent advances in data-driven intelligent control for wireless communication: A comprehensive survey," 2024.
- [9] W.-C. Tsai, C.-W. Chen, and A.-Y. Wu, "Deep unfolding-based channel estimation for irs-aided mmwave systems via two-stage lamp network with row compression," *IEEE Transactions on Vehicular Technology*, vol.73, no.11, pp.16832–16845, 2024.
- [10] R. Sun, N. Cheng, C. Li, W. Quan, H. Zhou, Y. Wang, W. Zhang, and X. Shen, "A comprehensive survey of knowledge-driven deep learning for intelligent wireless network optimization in 6g," *IEEE Communications Surveys and Tutorials*, pp.1–1, 2025.
- [11] G. Ma, X. Jiao, J. Mu, H. Han, and Y. Yang, "Deep learning-based detection for marker codes over insertion and deletion channels," *IEEE Transactions on Communications*, vol.72, no.10, pp.5945–5959, 2024.
- [12] X. He, H. Xu, J. Wang, W. Xie, X. Li, and A. Nallanathan, "Joint active and passive beamforming in ris-assisted covert symbiotic radio based on deep unfolding," *IEEE Transactions on Vehicular Technology*, vol.73, no.9, pp.14021–14026, 2024.
- [13] J. Zhang, M. Liu, J. Tang, N. Zhao, D. Niyato, and X. Wang, "Joint design for ris-aided isac via deep unfolding learning," *IEEE Transactions on Cognitive Communications and Networking*, vol.11, no.1, pp.349–361, 2025.
- [14] P. Krishnananthalingam, N. T. Nguyen, and M. Juntti, "Constant modulus waveform design for wideband multicarrier joint communications and sensing via deep unfolding," in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, pp.1–6, 2024.

# واژه‌نامه انگلیسی به فارسی

## E

Edge Device ..... دستگاه لبه

Energy Efficiency ..... کارایی انرژی

## G

Generative Artificial Intelligence .. هوش مصنوعی مولد

Graph Neural Network ..... شبکه عصبی گرافی

## H

Hallucination ..... توهم‌زایی

Hyperparameter ..... فرایارامتر

## I

Inference ..... استنتاج

Intelligent Reflecting Surface ... سطح بازتابنده هوشمند

Inter Symbol Interference ..... تداخل بین نمادها

Interpretability ..... تفسیر پذیری

## J

Joint Subcarrier and ... تخصیص مشترک زیر حامل و توان

## A

Algorithm Unfolding ..... گسترش الگوریتم

ante-hoc explainability ..... قابلیت تفسیر پیشینی

Attention Sub-Network ..... زیرشبکه توجه

## B

Beamfocusing ..... کانونی‌سازی پرتو

Beamforming ..... پرتوسازی

## C

Channel State Information ..... اطلاعات وضعیت کانال

Communication-specific ..... دانش دامنه خاص ارتباطات  
domain knowledge

Compressed Sensing ..... سنجش فشرده

Constraint-Specific Loss Function . تابع زیان مقید-محور

## D

Deep Unfolding ..... گسترش ژرف

Deep Learning ..... یادگیری ژرف

Deep Neural Network ..... شبکه عصبی ژرف

Domain Knowledge ..... دانش دامنه

Neural Sum Rate . . . . . بیشینه‌سازی نرخ جمع عصبی	Power Allocation
Maximization	
Nonconvex . . . . . غیرمحدب	
<b>K</b>	
	Knowledge-Driven Deep . . . . . یادگیری ژرف دانش‌محور
<b>P</b>	Learning
Pipeline . . . . . خط لوله	
post-hoc explainability . . . . . قابلیت تفسیر پسینی	
Precoder . . . . . پیش‌کدگذار	<b>L</b>
Primal-Dual Learning . . . . . یادگیری اولیه-دوگانه	Lagrange Duality . . . . . دوگانگی لاگرانژی
Projection Function . . . . . تابع افکنش	Large Language Model . . . . . مدل زبانی بزرگ
	Liquid Neural Networks . . . . . شبکه عصبی مایع
	Long Short-Term . . . . . حافظه طولانی مدت-کوتاه مدت
<b>Q</b>	Memory
Quality of Service . . . . . کیفیت خدمت	Loss Function . . . . . تابع زیان
<b>R</b>	
Real-Time . . . . . آنی	<b>M</b>
Recurrent Neural Network . . . . . شبکه عصبی بازگشتی	Majorization-Minimization . . . . . کاهش-بیشینه‌سازی
Refinement Structure . . . . . ساختار پالایش	Markov Decision Process . . . . . پردازش تصمیم‌گیری مارکوف
Reinforcement Learning . . . . . یادگیری تقویتی	Massive Millimeter Wave . . . . . موج میلی‌متری عظیم
	Metaverse . . . . . فراجهان
	Multi-Objective Learning . . . . . یادگیری چند-هدفه
	Mixing Structure . . . . . ساختار مختلط
	Millimeter Wave . . . . . موج میلی‌متری
<b>S</b>	Model-Driven . . . . . مدل-محور
Self-Interference Cancellation . . . . . لغو تداخل متوالی	
Substructure Design . . . . . طراحی زیرساخت	<b>N</b>
Standard Interference Function . . . . . تابع تداخل استاندارد	Near Field Communication . . . . . ارتباط میدان نزدیک

**T**

Trainable Parameter ..... پارامتر قابل آموزش



# واژه‌نامه فارسی به انگلیسی

Loss Function . . . . . تابع زیان	ا
Constraint-Specific Loss Function . . . . . تابع زیان مقید-محور	
Joint Subcarrier and . . . . . تخصیص مشترک زیر حامل و توان	Real-Time . . . . . آنی
Power Allocation . . . . . تخصیص توان	Near Field Communication . . . . . ارتباط میدان نزدیک
Inter Symbol Interference . . . . . تداخل بین نمادها	Inference . . . . . استنتاج
Interpretability . . . . . تفسیر پذیری	Channel State Information . . . . . اطلاعات وضعیت کانال
Hallucination . . . . . توهم‌زایی	

## ب

## ح

Neural Sum Rate . . . . . بیشینه‌سازی نرخ جمع عصبی	
Long Short-Term Memory . . . . . حافظه طولانی مدت-کوتاه مدت	Maximization

## خ

## پ

Pipeline . . . . . خط لوله	Trainable Parameter . . . . . پارامتر قابل آموزش
	Beamforming . . . . . پرتوسازی
	Markov Decision Process . . . . . پردازش تصمیم‌گیری مارکوف
د	Precoder . . . . . پیش‌کدگذار

Domain Knowledge . . . . . دانش دامنه
Communication-specific . . . . . دانش دامنه خاص ارتباطات
domain knowledge

## ت

Edge Device . . . . . دستگاه لبه	Projection Function . . . . . تابع افکنش
Lagrange Duality . . . . . دوگانگی لاگرانژی	Standard Interference Function . . . . . تابع تداخل استاندارد

## غ

Nonconvex ..... غیرمحدب

## ز

Attention Sub-Network ..... زیرشبکه توجه

## ف

Hyperparameter ..... فراپارامتر

Metaverse ..... فراجهان

## س

## ق

Refinement Structure ..... ساختار پالایش

Mixing Structure ..... ساختار مختلط

Intelligent Reflecting Surface ... سطح بازتابنده هوشمند

Compressed Sensing ..... سنجش فشرده

post-hoc explainability ..... قابلیت تفسیر پسینی

ante-hoc explainability ..... قابلیت تفسیر پیشینی

## ش

## ک

Energy Efficiency ..... کارایی انرژی

Beamfocusing ..... کانونی‌سازی پرتو

Majorization-Minimization ..... کاهش-بیشینه‌سازی

Quality of Service ..... کیفیت خدمت

Recurrent Neural Network ..... شبکه عصبی بازگشتی

Deep Neural Network ..... شبکه عصبی ژرف

Graph Neural Network ..... شبکه عصبی گرافی

Liquid Neural Networks ..... شبکه عصبی مایع

## گ

## ط

Algorithm Unfolding ..... گسترش الگوریتم

Deep Unfolding ..... گسترش ژرف

Substructure Design ..... طراحی زیرساخت

## ل

Self-Interference Cancellation . . . . . لغو تداخل متوالی

## م

Large Language Model . . . . . مدل زبانی بزرگ

Model-Driven . . . . . مدل-محور

Millimeter Wave . . . . . موج میلی‌متری

Massive Millimeter Wave . . . . . موج میلی‌متری عظیم

## ه

Generative Artificial Intelligence . . هوش مصنوعی مولد

## ی

Primal-Dual Learning . . . . . یادگیری اولیه-دوگانه

Reinforcement Learning . . . . . یادگیری تقویتی

Multi-Objective Learning . . . . . یادگیری چند-هدفه

Deep Learning . . . . . یادگیری ژرف

Knowledge-Driven Deep . . . . . یادگیری ژرف دانش‌محور

Learning

محیط، ۴

معماری

UE<sup>1</sup>، ۴

نکته، ۶

---

<sup>1</sup>User Equipment